

# Exponent Function for Source Coding with Side Information at the Decoder at Rates below the Rate Distortion Function

Yasutada Oohama

University of Electro-Communications, Tokyo, Japan

Email: oohama@uec.ac.jp

**Abstract**—We consider the rate distortion problem with side information at the decoder posed and investigated by Wyner and Ziv. The rate distortion function indicating the trade-off between the rate on the data compression and the quality of data obtained at the decoder was determined by Wyner and Ziv. In this paper, we study the error probability of decoding at rates below the rate distortion function. We evaluate the probability of decoding such that the estimation of source outputs by the decoder has a distortion not exceeding a prescribed distortion level. We prove that when the rate of the data compression is below the rate distortion function this probability goes to zero exponentially and derive an explicit lower bound of this exponent function. On the Wyner-Ziv source coding problem the strong converse coding theorem has not been established yet. We prove this as a simple corollary of our result.

## I. SOURCE CODING WITH SIDE INFORMATION AT THE DECODER

Let  $\mathcal{X}$  and  $\mathcal{Y}$  be finite sets and  $\{(X_t, Y_t)\}_{t=1}^{\infty}$  be a stationary discrete memoryless source. For each  $t = 1, 2, \dots$ , the random pair  $(X_t, Y_t)$  takes values in  $\mathcal{X} \times \mathcal{Y}$ , and has a probability distribution

$$p_{XY} = \{p_{XY}(x, y)\}_{(x, y) \in \mathcal{X} \times \mathcal{Y}}$$

We write  $n$  independent copies of  $\{X_t\}_{t=1}^{\infty}$  and  $\{Y_t\}_{t=1}^{\infty}$ , respectively as

$$X^n = X_1, X_2, \dots, X_n \text{ and } Y^n = Y_1, Y_2, \dots, Y_n.$$

We consider a communication system depicted in Fig. 1. Data sequences  $X^n$  is separately encoded to  $\varphi^{(n)}(X^n)$  and is sent to the information processing center. At the center the decoder function  $\psi^{(n)}$  observes  $\varphi^{(n)}(X^n)$  and  $Y^n$  to output the estimation  $Z^n$  of  $X^n$ . The encoder function  $\varphi^{(n)}$  is defined by

$$\varphi^{(n)} : \mathcal{X}^n \rightarrow \mathcal{M}_n = \{1, 2, \dots, M_n\}, \quad (1)$$

where  $\|\varphi^{(n)}\| (= M_n)$  stands for the range of cardinality of  $\varphi^{(n)}$ . Let  $\mathcal{Z}$  be a reproduction alphabet. The decoder function  $\psi^{(n)}$  is defined by

$$\psi^{(n)} : \mathcal{M}_n \times \mathcal{Y}^n \rightarrow \mathcal{Z}^n. \quad (2)$$

Let  $d : \mathcal{X} \times \mathcal{Z} \rightarrow [0, \infty)$  be an arbitrary distortion measure on  $\mathcal{X} \times \mathcal{Z}$ . The distortion between  $x^n \in \mathcal{X}^n$  and  $z^n \in \mathcal{Z}^n$  is defined by

$$d(x^n, z^n) \triangleq \sum_{t=1}^n d(x_t, z_t).$$

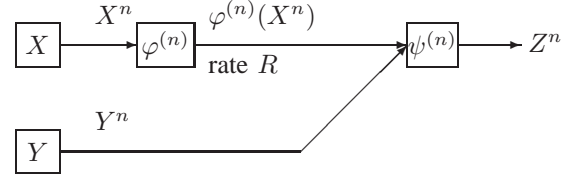


Fig. 1. Wyner-Ziv source coding system.

The excess-distortion probability of decoding is

$$P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) = \Pr \left\{ \frac{1}{n} d(X^n, Z^n) \geq \Delta \right\}, \quad (3)$$

where  $Z^n = \psi^{(n)}(\varphi^{(n)}(X^n), Y^n)$ . The average distortion  $\Delta^{(n)}$  between  $X^n$  and  $Z^n$  is defined by

$$\Delta^{(n)} \triangleq \frac{1}{n} E[d(X^n, Z^n)] \triangleq \frac{1}{n} \sum_{t=1}^n E[d(X_t, Z_t)].$$

A pair  $(R, \Delta)$  is  $\varepsilon$ -achievable for  $p_{XY}$  if there exist a sequence of pairs  $\{(\varphi^{(n)}, \psi^{(n)})\}_{n \geq 1}$  such that for any  $\delta > 0$  and any  $n$  with  $n \geq n_0 = n_0(\varepsilon, \delta)$ ,

$$\frac{1}{n} \log \|\varphi^{(n)}\| \leq R + \delta, \quad P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq \varepsilon.$$

The rate distortion region  $\mathcal{R}_{WZ}(\varepsilon | p_{XY})$  is defined by

$$\begin{aligned} & \mathcal{R}_{WZ}(\varepsilon | p_{XY}) \\ &= \{ (R, \Delta) : (R, \Delta) \text{ is } \varepsilon\text{-achievable for } p_{XY} \}. \end{aligned}$$

Furthermore set

$$\mathcal{R}_{WZ}(p_{XY}) \triangleq \bigcap_{\varepsilon > 0} \mathcal{R}_{WZ}(\varepsilon | p_{XY}).$$

On the other hand, we can define a rate distortion region based on the average distortion criterion, a formal definition of which is the following. A pair  $(R, \Delta)$  is *achievable* for  $p_{XY}$  if there exist a sequence of pairs  $\{(\varphi^{(n)}, \psi^{(n)})\}_{n \geq 1}$  such that for any  $\delta > 0$  and any  $n$  with  $n \geq n_0 = n_0(\delta)$ ,

$$\frac{1}{n} \log \|\varphi^{(n)}\| \leq R + \delta, \quad \Delta^{(n)} \leq \Delta + \delta.$$

The rate distortion region  $\tilde{\mathcal{R}}_{\text{WZ}}(p_{XY})$  is defined by

$$\begin{aligned} & \tilde{\mathcal{R}}_{\text{WZ}}(p_{XY}) \\ & \triangleq \{(R, \Delta) : (R, \Delta) \text{ is achievable for } p_{XY}\}. \end{aligned}$$

We can show that the three rate distortion regions  $\mathcal{R}_{\text{WZ}}(\varepsilon|p_{XY})$ ,  $\varepsilon \in (0, 1)$ ,  $\mathcal{R}_{\text{WZ}}(p_{XY})$ , and  $\tilde{\mathcal{R}}_{\text{WZ}}(p_{XY})$  satisfy the following property.

*Property 1:*

- a) The regions  $\mathcal{R}_{\text{WZ}}(\varepsilon|p_{XY})$ ,  $\varepsilon \in (0, 1)$ ,  $\mathcal{R}_{\text{WZ}}(p_{XY})$ , and  $\tilde{\mathcal{R}}_{\text{WZ}}(p_{XY})$  are closed convex sets of  $\mathbb{R}_+^2$ , where

$$\mathbb{R}_+^2 \triangleq \{(R, \Delta) : R \geq 0, \Delta \geq 0\}.$$

- b)  $\mathcal{R}_{\text{WZ}}(\varepsilon|p_{XY})$  has another form using  $(n, \varepsilon)$ -rate distortion region, the definition of which is as follows. We set

$$\begin{aligned} & \mathcal{R}_{\text{WZ}}(n, \varepsilon|p_{XY}) \\ & = \{(R, \Delta) : \text{There exists } (\varphi^{(n)}, \psi^{(n)}) \text{ such that} \\ & \quad \frac{1}{n} \log \|\varphi^{(n)}\| \leq R, \quad P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq \varepsilon\}, \end{aligned}$$

which is called the  $(n, \varepsilon)$ -rate distortion region. Using  $\mathcal{R}_{\text{WZ}}(n, \varepsilon|p_{XY})$ ,  $\mathcal{R}_{\text{WZ}}(\varepsilon|p_{XY})$  can be expressed as

$$\mathcal{R}_{\text{WZ}}(\varepsilon|p_{XY}) = \text{cl} \left( \bigcup_{m \geq 1} \bigcap_{n \geq m} \mathcal{R}_{\text{WZ}}(n, \varepsilon|p_{XY}) \right),$$

where  $\text{cl}(\cdot)$  stands for the closure operation.

Proof of this property is given in Appendix A. It is well known that  $\tilde{\mathcal{R}}_{\text{WZ}}(p_{XY})$  was determined by Wyner and Ziv [1]. To describe their result we introduce auxiliary random variables  $U$  and  $Z$ , respectively, taking values in finite sets  $\mathcal{U}$  and  $\mathcal{Z}$ . We assume that the joint distribution of  $(U, X, Y, Z)$  is

$$\begin{aligned} & p_{UXYZ}(u, x, y, z) \\ & = p_U(u)p_{X|U}(x|u)p_{Y|X}(y|x)p_{Z|UY}(z|u, y). \end{aligned}$$

The above condition is equivalent to

$$U \leftrightarrow X \leftrightarrow Y, X \leftrightarrow (U, Y) \leftrightarrow Z.$$

Define the set of probability distribution  $p = p_{UXYZ}$  by

$$\begin{aligned} \mathcal{P}(p_{XY}) & \triangleq \{p = p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}| + 1, \\ & \quad U \leftrightarrow X \leftrightarrow Y, X \leftrightarrow (U, Y) \leftrightarrow Z\}, \\ \mathcal{P}^*(p_{XY}) & \triangleq \{p = p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}| + 1, \\ & \quad U \leftrightarrow X \leftrightarrow Y, Z = \phi(U, Y) \\ & \quad \text{for some } \phi : \mathcal{U} \times \mathcal{Y} \rightarrow \mathcal{Z}\}. \end{aligned}$$

By definitions it is obvious that  $\mathcal{P}^*(p_{XY}) \subseteq \mathcal{P}(p_{XY})$ . Set

$$\begin{aligned} \mathcal{R}(p) & \triangleq \{(R, \Delta) : R, \Delta \geq 0, \\ & \quad R \geq I_p(X; U|Y), \Delta \geq E_p d(X, Z)\}, \\ \mathcal{R}(p_{XY}) & \triangleq \bigcup_{p \in \mathcal{P}(p_{XY})} \mathcal{R}(p), \\ \mathcal{R}^*(p_{XY}) & \triangleq \bigcup_{p \in \mathcal{P}^*(p_{XY})} \mathcal{R}(p). \end{aligned}$$

We can show that the above functions and sets satisfy the following property.

*Property 2:*

- a) The region  $\mathcal{R}(p_{XY})$  is a closed convex set of  $\mathbb{R}_+^2$ .  
b) For any  $p_{XY}$ , we have

$$\mathcal{R}(p_{XY}) = \mathcal{R}^*(p_{XY}).$$

Proof of Property 2 is given in Appendix C. In Property 2 part b),  $\mathcal{R}(p_{XY})$  is regarded as another expression of  $\mathcal{R}^*(p_{XY})$ . This expression is useful for deriving our main result. The rate region  $\mathcal{R}_{\text{WZ}}(p_{XY})$  was determined by Wyner and Ziv [1]. Their result is the following.

*Theorem 1 (Wyner and Ziv [1]):*

$$\tilde{\mathcal{R}}_{\text{WZ}}(p_{XY}) = \mathcal{R}^*(p_{XY}) = \mathcal{R}(p_{XY}).$$

On  $\mathcal{R}_{\text{WZ}}(p_{XY})$ , Csiszár and Körner [4] obtained the following result.

*Theorem 2 (Csiszár and Körner [4]):*

$$\mathcal{R}_{\text{WZ}}(p_{XY}) = \tilde{\mathcal{R}}_{\text{WZ}}(p_{XY}) = \mathcal{R}^*(p_{XY}) = \mathcal{R}(p_{XY}).$$

We are interested in an asymptotic behavior of the error probability of decoding to tend to one as  $n \rightarrow \infty$  for  $(R, \Delta) \notin \mathcal{R}_{\text{WZ}}(p_{XY})$ . To examine the rate of convergence, we define the following quantity. Set

$$\begin{aligned} & P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \triangleq 1 - P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta), \\ & G^{(n)}(R, \Delta|p_{XY}) \\ & \triangleq \min_{(\varphi^{(n)}, \psi^{(n)}) : \left(-\frac{1}{n}\right) \log P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq R} \left(-\frac{1}{n}\right) \log P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta). \end{aligned}$$

By time sharing we have that

$$\begin{aligned} & G^{(n+m)}\left(\frac{nR + mR'}{n+m}, \frac{n\Delta + m\Delta'}{n+m} \middle| p_{XY}\right) \\ & \leq \frac{nG^{(n)}(R, \Delta|p_{XY}) + mG^{(m)}(R', \Delta'|p_{XY})}{n+m}. \end{aligned} \quad (4)$$

Choosing  $R = R'$  and  $\Delta = \Delta'$  in (4), we obtain the following subadditivity property on  $\{G^{(n)}(R, \Delta|p_{XY})\}_{n \geq 1}$ :

$$\begin{aligned} & G^{(n+m)}(R, \Delta|p_{XY}) \\ & \leq \frac{nG^{(n)}(R, \Delta|p_{XY}) + mG^{(m)}(R, \Delta|p_{XY})}{n+m}, \end{aligned}$$

which together with Fekete's lemma yields that  $G^{(n)}(R, \Delta|p_{XY})$  exists and satisfies the following:

$$\lim_{n \rightarrow \infty} G^{(n)}(R, \Delta|p_{XY}) = \inf_{n \geq 1} G^{(n)}(R, \Delta|p_{XY}).$$

Set

$$\begin{aligned} & G(R, \Delta|p_{XY}) \triangleq \lim_{n \rightarrow \infty} G^{(n)}(R, \Delta|p_{XY}), \\ & \mathcal{G}(p_{XY}) \triangleq \{(R, \Delta, G) : G \geq G(R, \Delta|p_{XY})\}. \end{aligned}$$

The exponent function  $G(R, \Delta|p_{XY})$  is a convex function of  $(R, \Delta)$ . In fact, from (4), we have that for any  $\alpha \in [0, 1]$

$$\begin{aligned} & G(\alpha R + \bar{\alpha} R', \alpha \Delta + \bar{\alpha} \Delta'|p_{XY}) \\ & \leq \alpha G(R, \Delta|p_{XY}) + \bar{\alpha} G(R', \Delta'|p_{XY}). \end{aligned}$$

The region  $\mathcal{G}(p_{XY})$  is also a closed convex set. Our main aim is to find an explicit characterization of  $\mathcal{G}(p_{XY})$ . In this paper we derive an explicit outer bound of  $\mathcal{G}(p_{XY})$  whose section by the plane  $G = 0$  coincides with  $\mathcal{R}_{WZ}(p_{XY})$ .

## II. MAIN RESULT

In this section we state our main result. We first explain that the rate distortion region  $\mathcal{R}(p_{XY})$  can be expressed with two families of supporting hyperplanes. To describe this result we define two sets of probability distributions on  $\mathcal{U} \times \mathcal{X} \times \mathcal{Y}$  by

$$\begin{aligned}\mathcal{P}_{\text{sh}}(p_{XY}) &\triangleq \{p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}|, U \leftrightarrow X \leftrightarrow Y, \\ &\quad X \leftrightarrow (U, Y) \leftrightarrow Z\}, \\ \mathcal{Q} &\triangleq \{q = q_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|\}.\end{aligned}$$

We set

$$\begin{aligned}R^{(\mu)}(p_{XY}) &\triangleq \max_{p \in \mathcal{P}_{\text{sh}}(p_{XY})} \{\bar{\mu}I_p(X; U|Y) + \mu E_p d(X; Z)\},\end{aligned}$$

where  $\bar{\mu} = 1 - \mu$ . Furthermore set

$$\begin{aligned}\tilde{R}^{(\alpha, \mu)}(p_{XY}) &\triangleq \min_{q \in \mathcal{Q}} \{\bar{\alpha}[D(q_X||p_X) + D(q_{Y|XU}||p_{Y|X}q_{XU}) \\ &\quad + I_q(X; Z|UY)] \\ &\quad + 4\alpha[\bar{\mu}I_q(X; U|Y) + \mu E_q d(X, Z)]\}, \\ \mathcal{R}_{\text{sh}}(p_{XY}) &\triangleq \bigcap_{\mu \in [0, 1]} \{(R, \Delta) : \bar{\mu}R + \mu\Delta \geq R^{(\mu)}(p_{XY})\}, \\ \tilde{\mathcal{R}}_{\text{sh}}^{(\alpha)}(p_{XY}) &\triangleq \bigcap_{\mu \in [0, 1]} \left\{ (R, \Delta) : \bar{\mu}R + \mu\Delta \geq \frac{1}{4\alpha} \tilde{R}^{(\alpha, \mu)}(p_{XY}) \right\}, \\ \tilde{\mathcal{R}}_{\text{sh}}(p_{XY}) &\triangleq \bigcap_{\alpha \in (0, 1]} \tilde{\mathcal{R}}_{\text{sh}}^{(\alpha)}(p_{XY}) \\ &= \bigcap_{\substack{\mu \in [0, 1], \\ \alpha \in (0, 1]}} \left\{ (R, \Delta) : \bar{\mu}R + \mu\Delta \geq \frac{1}{4\alpha} \tilde{R}^{(\alpha, \mu)}(p_{XY}) \right\}.\end{aligned}$$

For  $\mathcal{R} \subseteq \mathbb{R}_+^2$ , we set

$$\mathcal{R} - \kappa(1, 1) \triangleq \{(a - \kappa, b - \kappa) \in \mathbb{R}_+^2 : (a, b) \in \mathcal{R}\}.$$

Then we have the following property.

*Property 3:* For any  $p_{XY}$ , we have

$$\mathcal{R}_{\text{sh}}(p_{XY}) = \mathcal{R}(p_{XY}). \quad (5)$$

For any  $\alpha \in (0, \alpha_0]$ , we have

$$\begin{aligned}R^{(\mu)}(p_{XY}) - c_1 \sqrt{\frac{\alpha}{\bar{\alpha}}} \log \left( c_2 \frac{\bar{\alpha}}{\alpha} \right) \\ \leq \frac{1}{4\alpha} \tilde{R}^{(\alpha, \mu)}(p_{XY}) \leq R^{(\mu)}(p_{XY}),\end{aligned} \quad (6)$$

where

$$\left. \begin{aligned}\alpha_0 &= \alpha_0(d_{\max}, |\mathcal{X}|) \triangleq [32 \log(|\mathcal{X}|e^{d_{\max}}) + 1]^{-1}, \\ c_1 &= c_1(d_{\max}, |\mathcal{X}|) \triangleq 4\sqrt{2 \log(|\mathcal{X}|e^{d_{\max}})}, \\ c_2 &= c_2(d_{\max}, |\mathcal{X}|, |\mathcal{Y}|, |\mathcal{Z}|) \triangleq \frac{e^{\frac{1}{2}d_{\max}} |\mathcal{Z}| |\mathcal{X}|^2 |\mathcal{Y}|^3}{8 \log(|\mathcal{X}|e^{d_{\max}})}.\end{aligned}\right\} \quad (7)$$

The two inequalities of (6) implies that for each  $\alpha \in (0, \alpha_0]$ ,

$$\begin{aligned}\mathcal{R}_{\text{sh}}(p_{XY}) - c_1 \sqrt{\frac{\alpha}{\bar{\alpha}}} \log \left( c_2 \frac{\bar{\alpha}}{\alpha} \right) (1, 1) \\ \subseteq \tilde{\mathcal{R}}_{\text{sh}}^{(\alpha)}(p_{XY}) \subseteq \mathcal{R}_{\text{sh}}(p_{XY}).\end{aligned}$$

Hence we have

$$\tilde{\mathcal{R}}_{\text{sh}}(p_{XY}) = \mathcal{R}_{\text{sh}}(p_{XY}). \quad (8)$$

Proof of Property 3 is given in Appendix D. For  $(\alpha, \mu) \in [0, 1]^2$ , define

$$\begin{aligned}\omega_{q||p}^{(\alpha, \mu)}(x, y, z|u) &\triangleq \bar{\alpha} \left[ \log \frac{q_X(x)}{p_X(x)} + \log \frac{q_{Y|XU}(y|x, u)}{p_{Y|X}(y|x)} \right. \\ &\quad \left. + \log \frac{q_{X|UYZ}(x|u, y, z)}{q_{X|UY}(x|u, y)} \right] \\ &\quad + 4\alpha \left[ \log \frac{q_{X|YU}(x|u, y)}{p_{X|Y}(x|y)} + \mu d(x, z) \right], \\ f_{q||p}^{(\alpha, \mu, \lambda)}(x, y, z|u) &\triangleq \exp \left\{ -\lambda \omega_{q||p}^{(\alpha, \mu)}(x, y, z|u) \right\}, \\ \Omega^{(\alpha, \mu, \lambda)}(q|p_{XY}) &\triangleq -\log E_q \left[ \exp \left\{ -\lambda \omega_{q||p}^{(\alpha, \mu)}(X, Y, Z|U) \right\} \right] \\ &= -\log \left[ \sum_{u, x, y, z} q(u, x, y, z) f_{q||p}^{(\alpha, \mu, \lambda)}(x, y, z|u) \right], \\ \Omega^{(\alpha, \mu, \lambda)}(p_{XY}) &\triangleq \min_{q \in \mathcal{Q}} \Omega^{(\alpha, \mu, \lambda)}(q|p_{XY}), \\ F^{(\alpha, \mu, \lambda)}(\bar{\mu}R + \mu\Delta|p_{XY}) &\triangleq \frac{\Omega^{(\alpha, \mu, \lambda)}(p_{XY}) - 4\alpha\lambda(\bar{\mu}R + \mu\Delta)}{1 + 4(1 - \alpha\mu)\lambda}, \\ F(R, \Delta|p_{XY}) &\triangleq \sup_{(\alpha, \mu) \in (0, 1]^2, \lambda > 0} F^{(\alpha, \mu, \lambda)}(\bar{\mu}R + \mu\Delta|p_{XY}), \\ \bar{\mathcal{G}}(p_{XY}) &\triangleq \{(R, \Delta, G) : G \geq F(R, \Delta|p_{XY})\}.\end{aligned}$$

We can show that the above functions and sets satisfy the following property.

*Property 4:*

- The cardinality bound  $|\mathcal{U}| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$  appearing in the definition of  $\Omega^{(\alpha, \mu, \lambda)}(p_{XY})$  is sufficient to describe this quantity.
- Define a probability distribution  $q^{(\lambda)} = q_{UXYZ}^{(\lambda)}$  by

$$\begin{aligned}q^{(\lambda)}(u, x, y, z) &\triangleq \frac{q(u, x, y, z) \exp \left\{ -\lambda \omega_{q||p}^{(\alpha, \mu)}(x, y, z|u) \right\}}{E_q \left[ \exp \left\{ -\lambda \omega_{q||p}^{(\alpha, \mu)}(X, Y, Z|U) \right\} \right]}.\end{aligned}$$

Then we have

$$\begin{aligned} \frac{d}{d\lambda} \Omega^{(\alpha, \mu, \lambda)}(q|p_{XY}) &= \mathbb{E}_{q^{(\lambda)}} \left[ \omega_{q|p}^{(\alpha, \mu)}(X, Y, Z|U) \right], \\ \frac{d^2}{d\lambda^2} \Omega^{(\alpha, \mu, \lambda)}(q|p_{XY}) &= -\text{Var}_{q^{(\lambda)}} \left[ \omega_{q|p}^{(\alpha, \mu)}(X, Y, Z|U) \right]. \end{aligned}$$

The second equality implies that  $\Omega_q^{(\alpha, \mu, \lambda)}(q|p_{XY})$  is a concave function of  $\lambda > 0$ .

c) Define

$$\begin{aligned} \rho &= \rho(p_{XY}) \\ &\triangleq \max_{q \in \mathcal{Q}} \max_{(\alpha, \mu) \in [0, 1]^2} \text{Var}_q \left[ \omega_{q|p}^{(\alpha, \mu)}(X, Y, Z|U) \right]. \end{aligned}$$

Since

$$0 \leq \left[ \omega_{q|p}^{(\alpha, \mu)}(x, y, z|u) \right]^2 < \infty$$

for  $(u, x, y, z) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ , we have  $\rho(p_{XY}) < \infty$ . Then for any  $\lambda \in (0, 1]$ , we have

$$\begin{aligned} &\Omega^{(\alpha, \mu, \lambda)}(q|p_{XY}) \\ &\geq \lambda \mathbb{E}_q \left[ \omega_{q|p}^{(\alpha, \mu)}(X, Y, Z|U) \right] - \frac{1}{2} \rho(p_{XY}) \lambda^2 \\ &= \bar{\alpha} [D(q_X||p_X) + D(q_{Y|XU}||p_{Y|X}|q_{XU}) \\ &\quad + I_q(X; Z|UY)] + 4\alpha [\bar{\mu} I_q(X; U|Y) \\ &\quad + \bar{\mu} D(q_{X|UY}||p_{X|Y}|q_{UY}) + \mu \mathbb{E}_q d(X, Z)] \\ &\quad - \frac{1}{2} \rho(p_{XY}) \lambda^2 \\ &\geq \bar{\alpha} [D(q_X||p_X) + D(q_{Y|XU}||p_{Y|X}|q_{XU}) \\ &\quad + I_q(X; Z|UY)] \\ &\quad + 4\alpha [\bar{\mu} I_q(X; U|Y) + \mu \mathbb{E}_q d(X, Z)] \\ &\quad - \frac{1}{2} \rho(p_{XY}) \lambda^2. \end{aligned}$$

Specifically, we have

$$\Omega^{(\alpha, \mu, \lambda)}(p_{XY}) \geq \lambda \tilde{R}^{(\alpha, \mu)}(p_{XY}) - \frac{1}{2} \rho(p_{XY}) \lambda^2.$$

d) For any  $\delta > 0$ , there exists a positive number  $\nu = \nu(\delta, d_{\max}, |\mathcal{X}|, |\mathcal{Y}|, |\mathcal{Z}|) \in (0, 1]$  such that for every  $\tau \in (0, \nu]$ , the condition  $(R + \tau, \Delta + \tau) \notin \mathcal{R}(p_{XY})$  implies

$$F(R, \Delta|p_{XY}) \geq \frac{\rho(p_{XY})}{2} \cdot g^2 \left( \frac{\tau^{3+\delta}}{\rho(p_{XY})} \right) > 0,$$

where  $g$  is the inverse function of  $\vartheta(a) \triangleq (1/2)a + a^2, a > 0$ .

Proof of Property 4 part a) is given in Appendix B. Proof of Property 4 parts b), c) and d) are given in Appendix E. Our main result is the following.

*Theorem 3:* For any  $R, \Delta > 0$ , any  $p_{XY}$ , and for any  $(\varphi^{(n)}, \psi^{(n)})$  satisfying

$$\frac{1}{n} \log M_n \leq R,$$

we have

$$\mathbb{P}_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq 5 \exp \{-nF(R, \Delta|p_{XY})\}. \quad (9)$$

It follows from Theorem 3 and Property 4 part d) that if  $(R, \Delta)$  is outside the rate distortion region, then the error probability of decoding goes to one exponentially and its exponent is not below  $F(R, \Delta|p_{XY})$ .

It immediately follows from Theorem 3 that we have the following corollary.

*Corollary 1:* For any  $R, \Delta > 0$  and any  $p_{XY}$ , we have

$$G(R, \Delta|p_{XY}) \geq F(R, \Delta|p_{XY}). \quad (10)$$

Furthermore for any  $p_{XY}$ , we have

$$\begin{aligned} \mathcal{G}(p_{XY}) &\subseteq \bar{\mathcal{G}}(p_{XY}) \\ &\triangleq \{(R, \Delta, G) : G \geq F(R, \Delta|p_{XY})\}. \end{aligned} \quad (11)$$

Proof of Theorem 3 will be given in the next section. The exponent function in the case of  $\Delta = 0$  can be obtained as a corollary of the result of Oohama and Han [3] for the separate source coding problem of correlated sources [2]. The techniques used by them is a method of types [4], which is not useful for proving Theorem 3. In fact when we use this method, it is very hard to extract a condition related to the Markov chain condition  $U \leftrightarrow X \leftrightarrow Y$ , which the auxiliary random variable  $U \in \mathcal{U}$  must satisfy when  $(R, \Delta)$  is on the boundary of the set  $\mathcal{R}(p_{XY})$ . Some novel techniques based on the information spectrum method introduced by Han [5] are necessary to prove this theorem.

From Theorem 3 and Property 4 part d), we obtain an explicit outer bound of  $\mathcal{R}_{WZ}(\varepsilon|p_{XY})$  with an asymptotically vanishing deviation from  $\mathcal{R}_{WZ}(p_{XY}) = \mathcal{R}(p_{XY})$ . The strong converse theorem immediately follows from this corollary. From Theorem 3 and Property 4 part d) we have the following corollary.

*Corollary 2:* For each fixed  $\varepsilon \in (0, 1)$  and for any  $\delta > 0$ , there exists a positive integer  $n_0$  with

$$n_0 = n_0(\varepsilon, \delta, d_{\max}, |\mathcal{X}|, |\mathcal{Y}|, |\mathcal{Z}|, \rho(p_{XY}))$$

such that for  $n \geq n_0$ , we have

$$\mathcal{R}_{WZ}(\varepsilon|p_{XY}) \subseteq \mathcal{R}(p_{XY}) - \kappa_n(1, 1),$$

where

$$\kappa_n \triangleq \left\{ \sqrt{\frac{\rho(p_{XY})}{2n}} \log \left( \frac{5}{1-\varepsilon} \right) + \frac{2}{n} \log \left( \frac{5}{1-\varepsilon} \right) \right\}^{\frac{1}{3+\delta}}.$$

It immediately follows from the above result that for each fixed  $\varepsilon \in (0, 1)$ , we have

$$\mathcal{R}_{WZ}(\varepsilon|p_{XY}) = \mathcal{R}_{WZ}(p_{XY}) = \mathcal{R}(p_{XY}).$$

Proof of this corollary will be given in the next section. The direct part of coding theorem, i.e., the inclusion of  $\mathcal{R}(p_{XY}) \subseteq \mathcal{R}_{WZ}(\varepsilon|p_{XY})$  was established by Csiszár and Körner [4]. They proved a weak converse theorem to obtain the inclusion  $\mathcal{R}_{WZ}(p_{XY}) \subseteq \mathcal{R}(p_{XY})$ . Until now we have had no result on the strong converse theorem. The above corollary stating the strong converse theorem for the Wyner-Ziv source coding problem implies that a long standing open problem since Csiszár and Körner [4] has been resolved.

### III. PROOF OF THE MAIN RESULT

In this section we prove Theorem 3. We first present a lemma which upper bounds the correct probability of decoding by the information spectrum quantities. We set

$$S_n \triangleq \varphi^{(n)}(X^n), Z^n \triangleq \psi_n(\varphi^{(n)}(X^n), Y^n).$$

It is obvious that

$$S_n \leftrightarrow X^n \leftrightarrow Y^n, X^n \leftrightarrow (S_n, Y^n) \leftrightarrow Z^n.$$

Then we have the following.

*Lemma 1:* For any  $\eta > 0$  and for any  $(\varphi^{(n)}, \psi^{(n)})$  satisfying

$$\frac{1}{n} \log M_n \leq R,$$

we have

$$P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq p_{S_n X^n Y^n Z^n} \left\{ \right.$$

$$\eta \geq \frac{1}{n} \log \frac{q_{X^n}^{(i)}(X^n)}{p_{X^n}(X^n)}, \quad (12)$$

$$\eta \geq \frac{1}{n} \log \frac{q_{Y^n|S_n X^n}^{(ii)}(Y^n|S_n, X^n)}{p_{Y^n|X^n}(Y^n|X^n)}, \quad (13)$$

$$\eta \geq \frac{1}{n} \log \frac{q_{X^n|S_n Y^n Z^n}^{(iii)}(X^n|S_n, Y^n, Z^n)}{p_{X^n|S_n Y^n}(X^n|S_n, Y^n)}, \quad (14)$$

$$R + \eta \geq \frac{1}{n} \log \frac{q_{X^n|S_n Y^n}^{(iv)}(X^n|S_n, Y^n)}{p_{X^n|Y^n}(X^n|Y^n)}, \quad (15)$$

$$\Delta \geq \frac{1}{n} \log \exp \{d(X^n, Z^n)\} + 4e^{-n\eta}. \quad (16)$$

The probability distribution and stochastic matrices appearing in the right members of (16) have a property that we can select them arbitrary. In (12), we can choose any probability distribution  $q_{X^n}^{(i)}$  on  $\mathcal{X}^n$ . In (13), we can choose any stochastic matrix  $q_{Y^n|S_n X^n}^{(ii)} : \mathcal{M}_n \times \mathcal{X}^n \rightarrow \mathcal{Y}^n$ . In (14), we can choose any stochastic matrix  $q_{X^n|S_n Y^n Z^n}^{(iii)} : \mathcal{M}_n \times \mathcal{Y}^n \times \mathcal{Z}^n \rightarrow \mathcal{X}^n$ . In (15), we can choose any stochastic matrix  $q_{X^n|S_n Y^n}^{(iv)} : \mathcal{M}_n \times \mathcal{Y}^n \rightarrow \mathcal{X}^n$ .

Proof of this lemma is given in Appendix F.

*Lemma 2:* Suppose that for each  $t = 1, 2, \dots, n$ , the joint distribution  $p_{S_n X^t Y^n}$  of the random vector  $S_n X^t Y^n$  is a marginal distribution of  $p_{S_n X^n Y^n}$ . Then, for  $t = 1, 2, \dots, n$ , we have the following Markov chain:

$$X_t \leftrightarrow S_n X^{t-1} Y^n \leftrightarrow Y^{t-1} \quad (17)$$

or equivalently that  $I(X_t; Y^{t-1} | S_n X^{t-1} Y^n) = 0$ .

Proof of this lemma is given in Appendix G. For  $t = 1, 2, \dots, n$ , set  $u_t \triangleq (s, x^{t-1}, y_{t+1}^n)$ . Let  $U_t \triangleq (S_n, X^{t-1}, Y_{t+1}^n)$  be a random vector taking values in  $\mathcal{M}_n \times \mathcal{X}^{t-1} \times \mathcal{Y}_{t+1}^n$ . From Lemmas 1 and 2, we have the following.

*Lemma 3:* For any  $\eta > 0$  and for any  $(\varphi^{(n)}, \psi^{(n)})$  satisfying

$$\frac{1}{n} \log M_n \leq R,$$

we have the following:

$$\begin{aligned} P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) &\leq p_{S_n X^n Y^n Z^n} \left\{ \right. \\ \eta &\geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{X_t}^{(i)}(X_t)}{p_{X_t}(X_t)}, \\ \eta &\geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{Y_t|U_t X_t}^{(ii)}(Y_t|U_t, X_t)}{p_{Y_t|X_t}(Y_t|X_t)}, \\ \eta &\geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{X_t|U_t Y_t Z_t}^{(iii)}(X_t|U_t, Y_t, Z_t)}{p_{X_t|U_t Y_t}(X_t|U_t, Y_t)}, \\ R + \eta &\geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{X_t|U_t Y_t}^{(iv)}(X_t|U_t, Y_t)}{p_{X_t|Y_t}(X_t|Y_t)}, \\ \Delta &\geq \frac{1}{n} \sum_{t=1}^n \log e^{d(X_t, Z_t)} \left. \right\} + 4e^{-n\eta}, \end{aligned} \quad (18)$$

where for each  $t = 1, 2, \dots, n$ , the following probability distribution and stochastic matrices:

$$q_{X_t}^{(i)}, q_{Y_t|U_t X_t}^{(ii)}, q_{X_t|U_t Y_t Z_t}^{(iii)}, \text{ and } q_{X_t|U_t Y_t}^{(iv)}$$

appearing in the first term in the right members of (18) have a property that we can choose their values arbitrary.

*Proof:* On the probability distributions appearing in the right members of (16), we take the following choices. In (12), we choose  $q_{X^n}^{(i)}$  so that

$$q_{X^n}^{(i)}(X^n) = \prod_{t=1}^n q_{X_t}^{(i)}(X_t). \quad (19)$$

In (13), we choose  $q_{Y^n|S_n X^n}^{(ii)}$  so that

$$\begin{aligned} &q_{Y^n|S_n X^n}^{(ii)}(Y^n|S_n, X^n) \\ &= \prod_{t=1}^n q_{Y_t|S_n X^t Y_{t+1}^n}^{(ii)}(Y_t|S_n, X^t, Y_{t+1}^n) \\ &= \prod_{t=1}^n q_{Y_t|X_t U_t}^{(ii)}(Y_t|U_t, X_t). \end{aligned} \quad (20)$$

In (14), we choose  $q_{X^n|S_n Y^n Z^n}^{(iii)}$  so that

$$\begin{aligned} &q_{X^n|S_n Y^n Z^n}^{(iii)}(X^n|S_n, Y^n, Z^n) \\ &= \prod_{t=1}^n q_{X_t|S_n X^{t-1} Y_t^n Z_t}^{(iii)}(X_t|S_n, X^{t-1}, Y_t^n, Z_t) \\ &= \prod_{t=1}^n q_{X_t|U_t Y_t Z_t}^{(iii)}(X_t|U_t, Y_t, Z_t). \end{aligned} \quad (21)$$



In (14), we note that

$$\begin{aligned}
& p_{X^n|S_n Y^n}(X^n|S_n, Y^n) \\
&= \prod_{t=1}^n p_{X_t|S_n X^{t-1} Y^n}(X_t|S_n, X^{t-1}, Y^n) \\
&\stackrel{(a)}{=} \prod_{t=1}^n p_{X_t|S_n X^{t-1} Y_t^n}(X_t|S_n, X^{t-1}, Y_t^n) \\
&= \prod_{t=1}^n p_{X_t|U_t Y_t}(X_t|U_t, Y_t). \tag{22}
\end{aligned}$$

Step (a) follows from Lemma 2. In (15), we choose  $q_{X^n|S_n Y^n}^{(iv)}$  so that

$$\begin{aligned}
& q_{X^n|S_n Y^n}^{(iv)}(X^n|S_n, Y^n) \\
&= \prod_{t=1}^n q_{X_t|S_n X^{t-1} Y_t^n}^{(iv)}(X_t|S_n, X^{t-1}, Y_t^n) \\
&= \prod_{t=1}^n q_{X_t|U_t Y_t}^{(iv)}(X_t|U_t, Y_t). \tag{23}
\end{aligned}$$

From Lemma 1 and (19)-(23), we have the bound (18) in Lemma 4.  $\blacksquare$

For each  $t = 1, 2, \dots, n$ , let  $\mathcal{Q}(\mathcal{U}_t \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$  be a set of all probability distributions  $q_{U_t X_t Y_t Z_t}$  on

$$\mathcal{U}_t \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} = \mathcal{M}_n \times \mathcal{X}^t \times \mathcal{Y}^{n-t+1} \times \mathcal{Z}$$

such that the support of the marginal distribution  $q_{X_t Y_t}$  is included in that of  $p_{XY}$ . For  $t = 1, 2, \dots, n$ , we simply write  $\mathcal{Q}_t = \mathcal{Q}(\mathcal{U}_t \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ . Similarly, for  $t = 1, 2, \dots, n$ , we simply write  $q_t = q_{U_t X_t Y_t Z_t} \in \mathcal{Q}_t$ . Set

$$\begin{aligned}
\mathcal{Q}^n &\triangleq \prod_{t=1}^n \mathcal{Q}_t = \prod_{t=1}^n \mathcal{Q}(\mathcal{U}_t \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}), \\
q^n &\triangleq \{q_t\}_{t=1}^n \in \mathcal{Q}^n.
\end{aligned}$$

From Lemma 3, we immediately obtain the following lemma.

*Lemma 4:* For any  $\eta > 0$ , for any  $(\varphi^{(n)}, \psi^{(n)})$  satisfying

$$\frac{1}{n} \log M_n \leq R,$$

and for any  $q^n \in \mathcal{Q}^n$ , we have the following:

$$\begin{aligned}
P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) &\leq p_{S_n X^n Y^n Z^n} \left\{ \right. \\
&\eta \geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{X_t}(X_t)}{p_{X_t}(X_t)}, \\
&\eta \geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{Y_t|U_t X_t}(Y_t|U_t, X_t)}{p_{Y_t|X_t}(Y_t|X_t)}, \\
&\eta \geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{X_t|U_t Y_t Z_t}(X_t|U_t, Y_t, Z_t)}{p_{X_t|U_t Y_t}(X_t|U_t, Y_t)}, \\
&R + \eta \geq \frac{1}{n} \sum_{t=1}^n \log \frac{q_{X_t|U_t Y_t}(X_t|U_t, Y_t)}{p_{X_t|Y_t}(X_t|Y_t)},
\end{aligned}$$

$$\Delta \geq \frac{1}{n} \sum_{t=1}^n \log e^{d(X_t, Z_t)} \Bigg\} + 4e^{-n\eta}, \tag{24}$$

where for each  $t = 1, 2, \dots, n$ , the following probability distribution and stochastic matrices:

$$q_{X_t}, q_{Y_t|U_t X_t}, q_{X_t|U_t Y_t Z_t}, \text{ and } q_{X_t|U_t Y_t}$$

appearing in the first term in the right members of (24) are chosen so that they are induced by the joint distribution  $q_t = q_{U_t X_t Y_t Z_t} \in \mathcal{Q}_t$ .

To evaluate an upper bound of (24) in Lemma 4. We use the following lemma, which is well known as the Cramér's bound in the large deviation principle.

*Lemma 5:* For any real valued random variable  $Z$  and any  $\theta > 0$ , we have

$$\Pr\{Z \geq a\} \leq \exp[-(\theta a - \log E[\exp(\theta Z)])].$$

Here we define a quantity which serves as an exponential upper bound of  $P_c^{(n)}(\varphi^{(n)}, \psi^{(n)})$ . Let  $\mathcal{P}^{(n)}(p_{XY})$  be a set of all probability distributions  $p_{S_n X^n Y^n Z^n}$  on  $\mathcal{M}_n \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$  having the form:

$$\begin{aligned}
& p_{S_n X^n Y^n Z^n}(s, x^n, y^n, z^n) \\
&= p_{S_n|X^n}(s|x^n) \left\{ \prod_{t=1}^n p_{X_t Y_t}(x_t, y_t) \right\} p_{Z^n|Y^n S_n}(z^n|y^n, s).
\end{aligned}$$

For simplicity of notation we use the notation  $p^{(n)}$  for  $p_{S_n X^n Y^n Z^n} \in \mathcal{P}^{(n)}(p_{XY})$ . We assume that  $p_{U_t X_t Y_t Z_t} = p_{S_n X^t Y_t^n Z_t}$  is a marginal distribution of  $p^{(n)}$ . For  $t = 1, 2, \dots, n$ , we simply write  $p_t = p_{U_t X_t Y_t Z_t}$ . For  $p^{(n)} \in \mathcal{P}^{(n)}(p_{XY})$  and  $q^n \in \mathcal{Q}^n$ , we define

$$\begin{aligned}
& \Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n|p_{XY}) \\
&\triangleq -\log E_{p^{(n)}} \left[ \prod_{t=1}^n \exp \left\{ -\theta \omega_{p_t|q_t}^{(\alpha, \mu)}(X_t, Y_t, Z_t|U_t) \right\} \right] \\
&= -\log E_{p^{(n)}} \left[ \left\{ \prod_{t=1}^n \frac{p_{X_t}^{\bar{\alpha}\theta}(X_t)}{q_{X_t}^{\bar{\alpha}\theta}(X_t)} \frac{p_{Y_t|X_t}^{\bar{\alpha}\theta}(Y_t|X_t)}{q_{Y_t|X_t U_t}^{\bar{\alpha}\theta}(Y_t|X_t, U_t)} \right\} \right. \\
&\quad \times \left\{ \prod_{t=1}^n \frac{p_{X_t|U_t Y_t}^{\bar{\alpha}\theta}(X_t|U_t, Y_t)}{q_{X_t|U_t Y_t Z_t}^{\bar{\alpha}\theta}(X_t|U_t, Y_t, Z_t)} \right\} \\
&\quad \times \left. \left\{ \prod_{t=1}^n \frac{p_{X_t|Y_t}^{4\alpha\theta}(X_t|Y_t)}{q_{X_t|Y_t U_t}^{4\alpha\theta}(X_t|U_t, Y_t)} e^{4\alpha\mu\theta d(X_t, Z_t)} \right\} \right],
\end{aligned}$$

where for each  $t = 1, 2, \dots, n$ , the following probability distribution and stochastic matrices:

$$q_{X_t}, q_{X_t|U_t Y_t}, q_{X_t|U_t Y_t Z_t}, q_{Y_t|X_t U_t}$$

appearing in the definition of  $\Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n|p_{XY})$  are chosen so that they are induced by the joint distribution  $q_t = q_{U_t X_t Y_t Z_t} \in \mathcal{Q}_t$ .

Here we give a remark on an essential difference between  $p^{(n)} \in \mathcal{P}^{(n)}(p_{XY})$  and  $q^n \in \mathcal{Q}^n$ . For the former the probability distributions  $p_t$ ,  $t = 1, 2, \dots, n$ , are consistent with  $p^{(n)}$ , since they are marginal distributions of  $p^{(n)}$ . On the other hand, for the latter,  $q^n$  is just a sequence of  $n$  probability

distributions. Hence, we may not have the consistency between the  $n$  elements  $q_t$ ,  $t = 1, 2, \dots, n$ , of  $q^n$ .

By Lemmas 4 and 5, we have the following proposition.

*Proposition 1:* For any  $\alpha, \mu, \theta > 0$ , any  $q^n \in \mathcal{Q}^n$ , and any  $(\varphi^{(n)}, \psi^{(n)})$  satisfying

$$\frac{1}{n} \log M_n \leq R,$$

we have

$$\begin{aligned} & P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \\ & \leq 5 \exp \left\{ -n [1 + (3 + \alpha)\theta]^{-1} \right. \\ & \quad \times \left. \left[ \frac{1}{n} \Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}) - 4\alpha\theta(\bar{\mu}R + \mu\Delta) \right] \right\}. \end{aligned}$$

*Proof:* By Lemma 4, for  $\alpha, \mu, \theta \geq 0$ , we have the following chain of inequalities:

$$\begin{aligned} P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) & \leq p_{S_n X^n Y^n Z^n} \left\{ \right. \\ & \quad \bar{\alpha}\eta \geq \frac{\bar{\alpha}}{n} \sum_{t=1}^n \log \frac{q_{X_t}(X_t)}{p_{X_t}(X_t)}, \\ & \quad \bar{\alpha}\eta \geq \frac{\bar{\alpha}}{n} \sum_{t=1}^n \log \frac{q_{Y_t|U_t X_t}(Y_t|U_t, X_t)}{p_{Y_t|X_t}(Y_t|X_t)}, \\ & \quad \bar{\alpha}\eta \geq \frac{\bar{\alpha}}{n} \sum_{t=1}^n \log \frac{q_{X_t|U_t Y_t Z_t}(X_t|U_t, Y_t, Z_t)}{p_{X_t|U_t Y_t}(X_t|U_t, Y_t)}, \\ & \quad 4\alpha\bar{\mu}(R + \eta) \geq \frac{4\alpha\bar{\mu}}{n} \sum_{t=1}^n \log \frac{q_{X_t|U_t Y_t}(X_t|U_t, Y_t)}{p_{X_t|Y_t}(X_t|Y_t)}, \\ & \quad 4\alpha\mu\Delta \geq \frac{4\alpha\mu}{n} \sum_{t=1}^n \log e^{d(X_t, Z_t)} \left. \right\} + 4e^{-n\eta} \\ & \leq p_{S_n X^n Y^n Z^n} \left\{ 4\alpha(\bar{\mu}R + \mu\Delta) + (3 + \alpha - 4\alpha\mu)\eta \right. \\ & \quad \left. \geq \frac{1}{n} \sum_{t=1}^n \omega_{q_t|p_t}^{(\alpha, \mu)}(X_t, Y_t, Z_t|U_t) \right\} + 4e^{-n\eta} \\ & = p_{S_n X^n Y^n Z^n} \left\{ -\frac{1}{n} \sum_{t=1}^n \theta \omega_{q_t|p_t}^{(\alpha, \mu)}(X_t, Y_t, Z_t|U_t) \right. \\ & \quad \left. \geq -\theta [4\alpha(\bar{\mu}R + \mu\Delta) + (3 + \alpha - 4\alpha\mu)\eta] \right\} + 4e^{-n\eta} \\ & \stackrel{(a)}{\leq} \exp \left[ n \left\{ 4\alpha\theta(\bar{\mu}R + \mu\Delta) + (3 + \alpha - 4\alpha\mu)\theta\eta \right. \right. \\ & \quad \left. \left. - \frac{1}{n} \Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}) \right\} \right] + 4e^{-n\eta}. \end{aligned} \quad (25)$$

Step (a) follows from Lemma 5. We choose  $\eta$  so that

$$\begin{aligned} -\eta & = 4\alpha\theta(\bar{\mu}R + \mu\Delta) + \theta(3 + \alpha - 4\alpha\mu)\eta \\ & \quad - \frac{1}{n} \Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}). \end{aligned} \quad (26)$$

Solving (26) with respect to  $\eta$ , we have

$$\eta = \frac{\frac{1}{n} \Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}) - 4\alpha\theta(\bar{\mu}R + \mu\Delta)}{1 + (3 + \alpha - 4\alpha\mu)\theta}.$$

For this choice of  $\eta$  and (25), we have

$$\begin{aligned} & P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq 5e^{-n\eta} \\ & = 5 \exp \left\{ -n [1 + (3 + \alpha - 4\alpha\mu)\theta]^{-1} \right. \\ & \quad \times \left. \left[ \frac{1}{n} \Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}) - 4\alpha\theta(\bar{\mu}R + \mu\Delta) \right] \right\}, \end{aligned}$$

completing the proof.  $\blacksquare$

Set

$$\begin{aligned} & \underline{\Omega}^{(\alpha, \mu, \theta)}(p_{XY}) \\ & \triangleq \inf_{n \geq 1} \min_{p^{(n)} \in \mathcal{P}^{(n)}(p_{XY})} \max_{q^n \in \mathcal{Q}^n} \frac{1}{n} \Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}). \end{aligned}$$

By Proposition 1 we have the following corollary.

*Corollary 3:* For any  $(\alpha, \mu) \in [0, 1]^2$ , for any  $\theta > 0$ , and for any  $(\varphi^{(n)}, \psi^{(n)})$  satisfying

$$\frac{1}{n} \log \|\varphi^{(n)}\| \leq R,$$

we have

$$\begin{aligned} & P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \\ & \leq 5 \exp \left\{ -n \left[ \frac{\underline{\Omega}^{(\alpha, \mu, \theta)}(p_{XY}) - 4\alpha\theta(\bar{\mu}R + \mu\Delta)}{1 + (3 + \alpha - 4\alpha\mu)\theta} \right] \right\}. \end{aligned}$$

We shall call  $\underline{\Omega}^{(\alpha, \mu, \theta)}(p_{XY})$  the communication potential. The above corollary implies that the analysis of  $\underline{\Omega}^{(\alpha, \mu, \theta)}(p_{XY})$  leads to an establishment of a strong converse theorem for Wyner-Ziv source coding problem. In the following argument we drive an explicit lower bound of  $\underline{\Omega}^{(\alpha, \mu, \theta)}(p_{XY})$ . We use a new technique we call *the recursive method*. The recursive method is a powerful tool to drive a single letterized exponent function for rates below the rate distortion function. This method is also applicable to prove the exponential strong converse theorems for other network information theory problems [6], [7], [8].

For each  $t = 1, 2, \dots, n$ , set

$$\begin{aligned} & f_{q_t|p_t}^{(\alpha, \mu, \theta)}(x_t, y_t, z_t|u_t) \\ & \triangleq \exp \left\{ -\theta \omega_{q_t|p_t}^{(\alpha, \mu)}(x_t, y_t, z_t|u_t) \right\} \\ & = \frac{p_{X_t}^{\bar{\alpha}\theta}(x_t)}{q_{X_t}^{\bar{\alpha}\theta}(x_t)} \frac{p_{Y_t|X_t}^{\bar{\alpha}\theta}(y_t|x_t)}{q_{Y_t|X_t U_t}^{\bar{\alpha}\theta}(y_t|x_t, u_t)} \frac{p_{X_t|U_t Y_t}^{\bar{\alpha}\theta}(x_t|u_t, y_t)}{q_{X_t|U_t Y_t Z_t}^{\bar{\alpha}\theta}(x_t|u_t, y_t, z_t)} \\ & \quad \times \frac{p_{X_t|Y_t}^{4\alpha\theta}(x_t|y_t)}{q_{X_t|Y_t U_t}^{4\alpha\theta}(x_t|u_t, y_t)} e^{4\alpha\mu\theta d(x_t, z_t)}. \end{aligned}$$

By definition we have

$$\begin{aligned} & \exp \left\{ -\Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}) \right\} \\ & = \sum_{s, y^n} p_{S_n Y^n}(s, y^n) \sum_{x^n, z^n} p_{X^n Z^n | S_n Y^n}(x^n, z^n | s, y^n) \\ & \quad \times \prod_{t=1}^n f_{q_t|p_t}^{(\alpha, \mu, \theta)}(x_t, y_t, z_t|u_t). \end{aligned} \quad (27)$$

For each  $t = 1, 2, \dots, n$ , we define the conditional probability distribution

$$p_{X^t Z^t | S_n Y^n}^{(\alpha, \mu, \theta; q^t)} \triangleq \left\{ p_{X^t Z^t | S_n Y^n}^{(\alpha, \mu, \theta; q^t)}(x^t, z^t | s, y^n) \right\}_{(x^t, z^t, s, y^n) \in \mathcal{X}^t \times \mathcal{Z}^t \times \mathcal{M}_n \times \mathcal{Y}^n}$$

by

$$\begin{aligned} & p_{X^t Z^t | S_n Y^n}^{(\alpha, \mu, \theta; q^t)}(x^t, z^t | s, y^n) \\ & \triangleq C_t^{-1}(s, y^n) p_{X^t Z^t | S_n Y^n}(x^t, z^t | s, y^n) \\ & \quad \times \prod_{i=1}^t f_{q_i | p_i}^{(\alpha, \mu, \theta)}(x_i, y_i, z_i | u_i) \end{aligned}$$

where

$$\begin{aligned} C_t(s, y^n) & \triangleq \sum_{x^t, z^t} p_{X^t Z^t | S_n Y^n}(x^t, z^t | s, y^n) \\ & \quad \times \prod_{i=1}^t f_{q_i | p_i}^{(\alpha, \mu, \theta)}(x_i, y_i, z_i | u_i) \end{aligned} \quad (28)$$

are constants for normalization. For  $t = 1, 2, \dots, n$ , define

$$\Phi_{t, q^t}^{(\alpha, \mu, \theta)}(s, y^n) \triangleq C_t(s, y^n) C_{t-1}^{-1}(s, y^n), \quad (29)$$

where we define  $C_0(s, y^n) = 1$  for  $(s, y^n) \in \mathcal{M}_n \times \mathcal{Y}^n$ . Then we have the following lemma.

**Lemma 6:** For each  $t = 1, 2, \dots, n$ , and for any  $(s, y^n, x^t, z^t) \in \mathcal{M}_n \times \mathcal{Y}^n \times \mathcal{X}^t \times \mathcal{Z}^t$ , we have

$$\begin{aligned} & p_{X^t Z^t | S_n Y^n}^{(\alpha, \mu, \theta; q^t)}(x^t, z^t | s, y^n) = (\Phi_{t, q^t}^{(\alpha, \mu, \theta)}(s, y^n))^{-1} \\ & \quad \times p_{X^{t-1} Y^{t-1} | S_n Y^n}(x^{t-1}, z^{t-1} | s, y^n) \\ & \quad \times p_{X_t Z_t | S_n X^{t-1} Y^n}(x_t, z_t | s, x^{t-1}, z^{t-1}, y^n) \\ & \quad \times f_{q_t | p_t}^{(\alpha, \mu, \theta)}(x_t, y_t, z_t | u_t). \end{aligned} \quad (30)$$

$$\begin{aligned} & \Phi_{t, q^t}^{(\alpha, \mu, \theta)}(s, y^n) \\ & = \sum_{x^t, z^t} p_{X^{t-1} Z^{t-1} | S_n Y^n}(x^{t-1}, z^{t-1} | s, y^n) \\ & \quad \times p_{X_t Z_t | S_n X^{t-1} Y^n}(x_t, z_t | s, x^{t-1}, z^{t-1}, y^n) \\ & \quad \times f_{q_t | p_t}^{(\alpha, \mu, \theta)}(x_t, y_t, z_t | u_t). \end{aligned} \quad (31)$$

Furthermore, we have

$$\begin{aligned} & \exp \left\{ -\Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}) \right\} \\ & = \sum_{s, y^n} p_{S_n Y^n}(s, y^n) \prod_{t=1}^n \Phi_{t, q^t}^{(\alpha, \mu, \theta)}(s, y^n). \end{aligned} \quad (32)$$

The equality (32) in Lemma 6 is obvious from (27), (28), and (29). Proofs of (30) and (31) in this lemma are given in Appendix H. Next we define a probability distribution of the random pair  $(S_n, Y^n)$  taking values in  $\mathcal{M}_n \times \mathcal{Y}^n$  by

$$\begin{aligned} & p_{S_n Y^n}^{(\alpha, \mu, \theta; q^t)}(s, y^n) \\ & = \tilde{C}_t^{-1} p_{S_n Y^n}(s, y^n) \prod_{i=1}^t \Phi_{i, q^i}^{(\alpha, \mu, \theta)}(s, y^n), \end{aligned} \quad (33)$$

where  $\tilde{C}_t$  is a constant for normalization given by

$$\tilde{C}_t = \sum_{s, y^n} p_{S_n Y^n}(s, y^n) \prod_{i=1}^t \Phi_{i, q^i}^{(\alpha, \mu, \theta)}(s, y^n).$$

For  $t = 1, 2, \dots, n$ , define

$$\Lambda_{t, q^t}^{(\alpha, \mu, \theta)} \triangleq \tilde{C}_t \tilde{C}_{t-1}^{-1}, \quad (34)$$

where we define  $\tilde{C}_0 = 1$ . Then we have the following.

**Lemma 7:**

$$\exp \left\{ -\Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}) \right\} = \prod_{t=1}^n \Lambda_{t, q^t}^{(\alpha, \mu, \theta)}, \quad (35)$$

$$\begin{aligned} \Lambda_{t, q^t}^{(\alpha, \mu, \theta)} & = \sum_{s, y^n} p_{S_n Y^n}^{(\alpha, \mu, \theta; q^{t-1})}(s, y^n) \Phi_{t, q^t}^{(\alpha, \mu, \theta)}(s, y^n) \\ & = \sum_{s, y^n} p_{S_n Y^n}^{(\alpha, \mu, \theta; q^{t-1})}(s, y^n) \\ & \quad \times \sum_{x^{t-1}, z^{t-1}} p_{X^{t-1} Z^{t-1} | S_n Y^n}(x^{t-1}, z^{t-1} | s, y^n) \\ & \quad \times p_{X_t Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) \\ & \quad \times f_{q_t | p_t}^{(\alpha, \mu, \theta)}(x_t, y_t, z_t | u_t). \end{aligned} \quad (36)$$

**Proof:** By the equality (32) in Lemma 6, we have

$$\begin{aligned} & \exp \left\{ -\Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}) \right\} \\ & = \tilde{C}_n = \prod_{t=1}^n \tilde{C}_t \tilde{C}_{t-1}^{-1} \stackrel{(a)}{=} \prod_{t=1}^n \Lambda_{t, q^t}^{(\alpha, \mu, \theta)}. \end{aligned} \quad (37)$$

Step (a) follows from the definition (34) of  $\Lambda_{t, q^t}^{(\alpha, \mu, \theta)}$ . We next prove (36) in Lemma 7. Multiplying  $\Lambda_{t, q^t}^{(\alpha, \mu, \theta)} = \tilde{C}_t / \tilde{C}_{t-1}$  to both sides of (33), we have

$$\begin{aligned} & \Lambda_{t, q^t}^{(\alpha, \mu, \theta)} p_{S_n Y^n}^{(\alpha, \mu, \theta; q^t)}(s, y^n) \\ & = \tilde{C}_{t-1}^{-1} p_{S_n Y^n}(s, y^n) \prod_{i=1}^t \Phi_{i, q^i}^{(\alpha, \mu, \theta)}(s, y^n), \\ & = p_{S_n Y^n}^{(\alpha, \mu, \theta; q^{t-1})}(s, y^n) \Phi_{t, q^t}^{(\alpha, \mu, \theta)}(s, y^n). \end{aligned} \quad (38)$$

Taking summations of (38) and (39) with respect to  $(s, y^n)$ , we have (36) in Lemma 7.  $\blacksquare$

The following proposition is a mathematical core to prove our main result.

**Proposition 2:** For  $\theta \in (0, 1/\bar{\alpha})$ , we choose the parameter  $\lambda$  such that

$$\lambda = \frac{\theta}{1 - \bar{\alpha}\theta} \Leftrightarrow \theta = \frac{\lambda}{1 + \bar{\alpha}\lambda}. \quad (40)$$

Then for any  $\alpha, \mu > 0$  and for any  $\theta \in (0, 1/\bar{\alpha})$ , we have

$$\underline{\Omega}^{(\alpha, \mu, \theta)}(p_{XY}) \geq \frac{1}{1 + \bar{\alpha}\lambda} \Omega^{(\alpha, \mu, \lambda)}(p_{XY}). \quad (41)$$

**Proof:** Set

$$\begin{aligned} \hat{\mathcal{Q}}_n & \triangleq \{q = q_{UXYZ} : |\mathcal{U}| \leq |\mathcal{M}_n| |\mathcal{X}^{n-1}| |\mathcal{Y}^{n-1}|\}, \\ \hat{\Omega}_n^{(\alpha, \mu, \lambda)}(p_{XY}) & \triangleq \min_{q \in \hat{\mathcal{Q}}_n} \Omega^{(\alpha, \mu, \lambda)}(q | p_{XY}). \end{aligned}$$



Set

$$\begin{aligned}
& p_{S_n X^t Y^n Z_t}^{(\alpha, \mu, \theta; q^{t-1})}(s, x^t, y^n, z_t) \\
&= p_{U_t X_t Y_t Z_t}^{(\alpha, \mu, \theta; q^{t-1})}(u_t, x_t, y_t, z_t) \\
&\triangleq \sum_{y^{t-1}, z^{t-1}} p_{S_n Y^n}^{(\alpha, \mu, \theta; q^{t-1})}(s, y^n) \\
&\quad \times p_{X^{t-1} Z^{t-1} | S_n Y^n}(x^{t-1}, z^{t-1} | s, y^n) \\
&\quad \times p_{X_t Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n). \quad (42)
\end{aligned}$$

Then by Lemma 7, we have

$$\begin{aligned}
\Lambda_{t, q^t}^{(\alpha, \mu, \theta)} &= \sum_{u_t, x_t, y_t, z_t} p_{U_t X_t Y_t Z_t}^{(\alpha, \mu, \theta; q^{t-1})}(u_t, x_t, y_t, z_t) \\
&\quad \times f_{q_t | p_t}^{(\alpha, \mu, \theta)}(x_t, y_t, z_t | u_t).
\end{aligned}$$

For each  $t = 1, 2, \dots, n$ , we recursively choose  $q_t = q_{U_t X_t Y_t Z_t}$  so that

$$q_{U_t X_t Y_t Z_t}(u_t, x_t, y_t, z_t) = p_{U_t X_t Y_t Z_t}^{(\alpha, \mu, \theta; q^{t-1})}(u_t, x_t, y_t, z_t)$$

and choose  $q_{X_t}$ ,  $q_{Y_t | X_t U_t}$ ,  $q_{X_t | U_t Y_t Z_t}$ , and  $q_{X_t | Y_t U_t}$  appearing in

$$\begin{aligned}
f_{q_t | p_t}^{(\alpha, \mu, \theta)}(x_t, y_t, z_t | u_t) &= \frac{p_{X_t}^{\bar{\alpha}\theta}(x_t)}{q_{X_t}^{\bar{\alpha}\theta}(x_t)} \frac{p_{Y_t | X_t}^{\bar{\alpha}\theta}(y_t | x_t)}{q_{Y_t | X_t U_t}^{\bar{\alpha}\theta}(y_t | x_t, u_t)} \\
&\times \frac{p_{X_t | U_t Y_t}^{\bar{\alpha}\theta}(x_t | u_t, y_t)}{q_{X_t | U_t Y_t Z_t}^{\bar{\alpha}\theta}(x_t | u_t, y_t, z_t)} \frac{p_{X_t | Y_t}^{4\alpha\theta}(x_t | y_t)}{q_{X_t | Y_t U_t}^{4\alpha\theta}(x_t | u_t, y_t)} e^{4\alpha\mu\theta d(x_t, z_t)}
\end{aligned}$$

such that they are the distributions induced by  $q_{U_t X_t Y_t Z_t}$ . Then for each  $t = 1, 2, \dots, n$ , we have the following chain of inequalities:

$$\begin{aligned}
& \Lambda_{t, q^t}^{(\alpha, \mu, \theta)} \\
&= \mathbb{E}_{q_t} \left[ \left\{ \frac{p_{X_t}^{\bar{\alpha}\theta}(X_t)}{q_{X_t}^{\bar{\alpha}\theta}(X_t)} \frac{p_{Y_t | X_t}^{\bar{\alpha}\theta}(Y_t | X_t)}{q_{Y_t | X_t U_t}^{\bar{\alpha}\theta}(Y_t | X_t, U_t)} \right\} \right. \\
&\quad \times \left\{ \frac{p_{X_t | U_t Y_t}^{\bar{\alpha}\theta}(X_t | U_t, Y_t)}{q_{X_t | U_t Y_t Z_t}^{\bar{\alpha}\theta}(X_t | U_t, Y_t, Z_t)} \right\} \\
&\quad \times \left\{ \frac{p_{X_t | Y_t}^{4\alpha\bar{\mu}\theta}(X_t | Y_t)}{q_{X_t | Y_t U_t}^{4\alpha\bar{\mu}\theta}(X_t | U_t, Y_t)} e^{4\alpha\mu\theta d(X_t, Z_t)} \right\} \Bigg] \\
&= \mathbb{E}_{q_t} \left[ \left\{ \frac{p_{X_t}^{\bar{\alpha}}(X_t)}{q_{X_t}^{\bar{\alpha}}(X_t)} \frac{p_{Y_t | X_t}^{\bar{\alpha}}(Y_t | X_t)}{q_{Y_t | X_t U_t}^{\bar{\alpha}}(Y_t | X_t, U_t)} \right\}^\theta \right. \\
&\quad \times \left\{ \frac{q_{X_t | U_t Y_t}^{\bar{\alpha}}(X_t | U_t, Y_t)}{q_{X_t | U_t Y_t Z_t}^{\bar{\alpha}}(X_t | U_t, Y_t, Z_t)} \right\}^\theta \\
&\quad \times \left\{ \frac{p_{X_t | Y_t}^{4\alpha\bar{\mu}}(X_t | Y_t)}{q_{X_t | Y_t U_t}^{4\alpha\bar{\mu}}(X_t | U_t, Y_t)} e^{4\alpha\mu d(X_t, Z_t)} \right\}^\theta \\
&\quad \times \left. \left\{ \frac{p_{X_t | U_t Y_t}(X_t | U_t, Y_t)}{q_{X_t | U_t Y_t}(X_t | U_t, Y_t)} \right\}^{\bar{\alpha}\theta} \right] \\
&\stackrel{(a)}{\leq} \left( \mathbb{E}_{q_t} \left[ \left\{ \frac{p_{X_t}^{\bar{\alpha}}(X_t)}{q_{X_t}^{\bar{\alpha}}(X_t)} \frac{p_{Y_t | X_t}^{\bar{\alpha}}(Y_t | X_t)}{q_{Y_t | X_t U_t}^{\bar{\alpha}}(Y_t | X_t, U_t)} \right\}^{\frac{\theta}{1-\bar{\alpha}\theta}} \right. \right.
\end{aligned}$$

$$\begin{aligned}
& \times \left. \left\{ \frac{q_{X_t | U_t Y_t}^{\bar{\alpha}}(X_t | U_t, Y_t)}{q_{X_t | U_t Y_t Z_t}^{\bar{\alpha}}(X_t | U_t, Y_t, Z_t)} \right\}^{\frac{\theta}{1-\bar{\alpha}\theta}} \right]^{1-\bar{\alpha}\theta} \\
&\times \left\{ \frac{p_{X_t | Y_t}^{4\alpha\bar{\mu}}(X_t | Y_t)}{q_{X_t | Y_t U_t}^{4\alpha\bar{\mu}}(X_t | U_t, Y_t)} e^{4\alpha\mu d(X_t, Z_t)} \right\}^{\frac{\theta}{1-\bar{\alpha}\theta}} \Bigg]^{1-\bar{\alpha}\theta} \\
&\times \left( \mathbb{E}_{q_t} \left\{ \frac{p_{X_t | U_t Y_t}(X_t | U_t, Y_t)}{q_{X_t | U_t Y_t}(X_t | U_t, Y_t)} \right\}^{\bar{\alpha}\theta} \right) \\
&= \exp \left\{ -(1-\bar{\alpha}\theta) \Omega^{(\alpha, \mu, \frac{\theta}{1-\bar{\alpha}\theta})}(q_t | p_{XY}) \right\} \\
&\stackrel{(b)}{=} \exp \left\{ -\frac{1}{1+\bar{\alpha}\lambda} \Omega^{(\alpha, \mu, \lambda)}(q_t | p_{XY}) \right\} \\
&\stackrel{(c)}{\leq} \exp \left\{ -\frac{1}{1+\bar{\alpha}\lambda} \hat{\Omega}_n^{(\alpha, \mu, \lambda)}(p_{XY}) \right\} \\
&\stackrel{(d)}{=} \exp \left\{ -\frac{1}{1+\bar{\alpha}\lambda} \Omega^{(\alpha, \mu, \lambda)}(p_{XY}) \right\}. \quad (43)
\end{aligned}$$

Step (a) follows from Hölder's inequality. Step (b) follows from (40). Step (c) follows from the definition of  $\hat{\Omega}_n^{(\alpha, \mu, \lambda)}(p_{XY})$ . Step (d) follows from that by Property 4 part a), the bound  $|\mathcal{U}| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$  is sufficient to describe  $\hat{\Omega}_n^{(\alpha, \mu, \lambda)}(p_{XY})$ . Hence, we have the following:

$$\begin{aligned}
& \max_{q^n \in \mathcal{Q}^n} \frac{1}{n} \Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}) \\
&\geq \frac{1}{n} \Omega^{(\alpha, \mu, \theta)}(p^{(n)}, q^n | p_{XY}) \stackrel{(a)}{=} -\frac{1}{n} \sum_{t=1}^n \log \Lambda_{t, q^t}^{(\alpha, \mu, \theta)} \\
&\stackrel{(b)}{\geq} \frac{1}{1+\bar{\alpha}\lambda} \Omega^{(\alpha, \mu, \lambda)}(p_{XY}). \quad (44)
\end{aligned}$$

Step (a) follows from (35) in Lemma 7. Step (b) follows from (43). Since (44) holds for any  $n \geq 1$  and any  $p^{(n)} \in \mathcal{P}^{(n)}(p_{XY})$ , we have

$$\underline{\Omega}^{(\alpha, \mu, \theta)}(p_{XY}) \geq \frac{1}{1+\bar{\alpha}\lambda} \Omega^{(\alpha, \mu, \lambda)}(p_{XY}).$$

Thus we have (41) in Proposition 2. ■

*Proof of Theorem 3:* For  $\theta \in (0, 1/\bar{\alpha})$ , set

$$\lambda = \frac{\theta}{1-\bar{\alpha}\theta} \Leftrightarrow \theta = \frac{\lambda}{1+\bar{\alpha}\lambda}. \quad (45)$$

Then we have the following:

$$\begin{aligned}
& \frac{1}{n} \log \left\{ \frac{5}{P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta)} \right\} \\
&\stackrel{(a)}{\geq} \frac{\underline{\Omega}^{(\alpha, \mu, \theta)}(p_{XY}) - 4\alpha\theta(\bar{\mu}R + \mu\Delta)}{1 + \theta(3 + \alpha - 4\alpha\mu)} \\
&\stackrel{(b)}{\geq} \frac{\frac{1}{1+\bar{\alpha}\lambda} \Omega^{(\alpha, \mu, \lambda)}(p_{XY}) - \frac{4\alpha\lambda}{1+\bar{\alpha}\lambda}(\bar{\mu}R + \mu\Delta)}{1 + \frac{\lambda}{1+\bar{\alpha}\lambda}(3 + \alpha - 4\alpha\mu)} \\
&= \frac{\Omega^{(\alpha, \mu, \lambda)}(p_{XY}) - 4\alpha\lambda(\bar{\mu}R + \mu\Delta)}{1 + \bar{\alpha}\lambda + \lambda(3 + \alpha - 4\alpha\mu)} \\
&= F^{(\alpha, \mu, \lambda)}(\bar{\mu}R + \mu\Delta | p_{XY}).
\end{aligned}$$

Step (a) follows from Corollary 3. Step (b) follows from Proposition 2 and (45). Since the above bound holds for any

positive  $\alpha$ ,  $\mu$ , and  $\lambda$ , we have

$$\frac{1}{n} \log \left\{ \frac{5}{P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta)} \right\} \geq F(R, \Delta | p_{XY}).$$

Thus (9) in Theorem 3 is proved.  $\blacksquare$

*Proof of Corollary 2:* To prove this corollary we use the following expression of  $\mathcal{R}_{WZ}(\varepsilon | p_{XY})$  stated in Property 1 part b):

$$\mathcal{R}_{WZ}(\varepsilon | p_{XY}) = \text{cl} \left( \bigcup_{m \geq 1} \bigcap_{n \geq m} \mathcal{R}_{WZ}(n, \varepsilon | p_{XY}) \right). \quad (46)$$

We assume that

$$(R, \Delta) \in \bigcup_{m \geq 1} \bigcap_{n \geq m} \mathcal{R}_{WZ}(n, \varepsilon | p_{XY}). \quad (47)$$

Then there exists a positive integer  $m_0 = m_0(\varepsilon)$  and some  $\{(\varphi^{(n)}, \psi^{(n)})\}_{n \geq m_0}$  such that for  $n \geq m_0(\varepsilon)$ , we have

$$\frac{1}{n} \log \|\varphi^{(n)}\| \leq R, \quad P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq \varepsilon.$$

Then by Theorem 3, we have

$$\begin{aligned} 1 - \varepsilon &\leq P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \\ &\leq 5 \exp \{-nF(R, \Delta | p_{XY})\} \end{aligned} \quad (48)$$

for any  $n \geq m_0(\varepsilon)$ . Fix any  $\delta > 0$ . We take a positive number  $\nu = \nu(\delta, d_{\max}, |\mathcal{X}|, |\mathcal{Y}|, |\mathcal{Z}|) \in (0, 1]$  appearing in Property 4 part d) and set

$$\begin{aligned} \kappa_n &= \left\{ \rho(p_{XY}) \vartheta \left( \sqrt{\frac{2}{n\rho(p_{XY})} \log \left( \frac{5}{1-\varepsilon} \right)} \right) \right\}^{\frac{1}{3+\delta}} \\ &= \left\{ \sqrt{\frac{\rho(p_{XY})}{2n} \log \left( \frac{5}{1-\varepsilon} \right)} + \frac{2}{n} \log \left( \frac{5}{1-\varepsilon} \right) \right\}^{\frac{1}{3+\delta}}. \end{aligned} \quad (49)$$

Since  $g$  is an inverse function of  $\vartheta$ , (49) is equivalent to

$$g \left( \frac{\kappa_n^{3+\delta}}{\rho(p_{XY})} \right) = \sqrt{\frac{2}{n\rho(p_{XY})} \log \left( \frac{5}{1-\varepsilon} \right)}. \quad (50)$$

We take a sufficiently large positive integer  $m_1$  so that we have  $\kappa_n < \nu$  for  $n \geq m_1$ . Set  $n_0 = \max\{m_0, m_1\}$ . We claim that for  $n \geq n_0$ , we have  $(R + \kappa_n, \Delta + \kappa_n) \in \mathcal{R}(p_{XY})$ . To prove this claim we suppose that  $(R + \kappa_{n^*}, \Delta + \kappa_{n^*})$  does not belong to  $\mathcal{R}(p_{XY})$  for some  $n^* \geq n_0$ . Since  $\mathcal{R}(p_{XY})$  is a closed set, there exists a positive number  $\tau > \kappa_{n^*}$  sufficiently close to  $\kappa_{n^*}$  such that

$$\begin{aligned} \kappa_{n^*} &< \tau \leq \nu, \\ (R + \tau, \Delta + \tau) &\notin \mathcal{R}(p_{XY}). \end{aligned}$$

Then we have the following chain of inequalities:

$$\begin{aligned} &5 \exp[-n^* F(R, \Delta | p_{XY})] \\ &\stackrel{(a)}{\leq} 5 \exp \left[ -\frac{n^* \rho(p_{XY})}{2} \cdot g^2 \left( \frac{\tau^{3+\delta}}{\rho(p_{XY})} \right) \right] \\ &\stackrel{(b)}{<} 5 \exp \left[ -\frac{n^* \rho(p_{XY})}{2} \cdot g^2 \left( \frac{\kappa_{n^*}^{3+\delta}}{\rho(p_{XY})} \right) \right] \\ &\stackrel{(c)}{=} 5 \exp \left[ -\frac{n^* \rho(p_{XY})}{2} \frac{2}{n^* \rho(p_{XY})} \log \left( \frac{5}{1-\varepsilon} \right) \right] \\ &= 1 - \varepsilon. \end{aligned} \quad (51)$$

Step (a) follows from Property 4 part d). Step (b) follows from  $\kappa_{n^*} < \tau$ . Step (c) follows from (50). The bound (51) contradicts (48). Hence we have  $(R + \kappa_n, \Delta + \kappa_n) \in \mathcal{R}(p_{XY})$  or equivalent to

$$(R, \Delta) \in \mathcal{R}(p_{XY}) - \kappa_n(1, 1)$$

for  $n \geq n_0$ . Recalling the first assumption (47) on  $(R, \Delta)$ , we obtain

$$\bigcup_{m \geq 1} \bigcap_{n \geq m} \mathcal{R}_{WZ}(n, \varepsilon | p_{XY}) \subseteq \mathcal{R}(p_{XY}) - \kappa_n(1, 1). \quad (52)$$

Taking the closure of both sides of (52), using (46), and considering that  $\mathcal{R}(p_{XY}) - \kappa_n(1, 1)$  is a closed set, we have that for  $n \geq n_0$ ,

$$\begin{aligned} &\mathcal{R}_{WZ}(\varepsilon | p_{XY}) \\ &= \text{cl} \left( \bigcup_{m \geq 1} \bigcap_{n \geq m} \mathcal{R}_{WZ}(n, \varepsilon | p_{XY}) \right) \subseteq \mathcal{R}(p_{XY}) - \kappa_n(1, 1), \end{aligned}$$

completing the proof.  $\blacksquare$

## APPENDIX

### A. Properties of the Rate Distortion Regions

In this appendix we prove Property 1. Property 1 part a) can easily be proved by the definitions of the rate distortion regions. We omit the proofs of this part. In the following argument we prove part b).

*Proof of Property 1 part b:* We set

$$\underline{\mathcal{R}}_{WZ}(m, \varepsilon | p_{XY}) = \bigcap_{n \geq m} \mathcal{R}_{WZ}(n, \varepsilon | p_{XY}).$$

By the definitions of  $\underline{\mathcal{R}}_{WZ}(m, \varepsilon | p_{XY})$  and  $\mathcal{R}_{WZ}(\varepsilon | p_{XY})$ , we have that  $\underline{\mathcal{R}}_{WZ}(m, \varepsilon | p_{XY}) \subseteq \mathcal{R}_{WZ}(\varepsilon | p_{XY})$  for  $m \geq 1$ . Hence we have that

$$\bigcup_{m \geq 1} \underline{\mathcal{R}}_{WZ}(m, \varepsilon | p_{XY}) \subseteq \mathcal{R}_{WZ}(\varepsilon | p_{XY}). \quad (53)$$

We next assume that  $(R, \Delta) \in \mathcal{R}_{WZ}(\varepsilon | p_{XY})$ . Set

$$\mathcal{R}_{WZ}^{(\delta)}(\varepsilon | p_{XY}) \triangleq \{(R + \delta, \Delta) : (R, \Delta) \in \mathcal{R}_{WZ}(\varepsilon | p_{XY})\}$$

Then, by the definitions of  $\mathcal{R}_{WZ}(n, \varepsilon | p_{XY})$  and  $\mathcal{R}_{WZ}(\varepsilon | p_{XY})$ , we have that for any  $\delta > 0$ , there exists  $n_0(\delta)$  such that for any  $n \geq n_0(\delta)$ ,

$$(R + \delta, \Delta) \in \mathcal{R}_{WZ}(n, \varepsilon | p_{XY}),$$

which implies that

$$\begin{aligned} \mathcal{R}_{\text{WZ}}^{(\delta)}(\varepsilon|p_{XY}) &\subseteq \bigcup_{n \geq n_0(\delta)} \mathcal{R}_{\text{WZ}}(n, \varepsilon|p_{XY}) \\ &= \underline{\mathcal{R}}_{\text{WZ}}(n_0(\delta), \varepsilon|p_{XY}) \\ &\subseteq \text{cl} \left( \bigcup_{m \geq 1} \underline{\mathcal{R}}_{\text{WZ}}(m, \varepsilon|p_{XY}) \right). \end{aligned} \quad (54)$$

Here we assume that there exists a pair  $(R, \Delta)$  belonging to  $\mathcal{R}_{\text{WZ}}(\varepsilon|p_{XY})$  such that

$$(R, \Delta) \notin \text{cl} \left( \bigcup_{m \geq 1} \underline{\mathcal{R}}_{\text{WZ}}(m, \varepsilon|p_{XY}) \right). \quad (55)$$

Since the set in the right hand side of (55) is a closed set, we have

$$(R + \delta, \Delta) \notin \text{cl} \left( \bigcup_{m \geq 1} \underline{\mathcal{R}}_{\text{WZ}}(m, \varepsilon|p_{XY}) \right) \quad (56)$$

for some small  $\delta > 0$ . Note that  $(R + \delta, \Delta) \in \mathcal{R}_{\text{WZ}}^{(\delta)}(\varepsilon|p_{XY})$ . Then (56) contradicts (54). Thus we have

$$\begin{aligned} &\bigcup_{m \geq 1} \underline{\mathcal{R}}_{\text{WZ}}(m, \varepsilon|p_{XY}) \\ &\subseteq \mathcal{R}_{\text{WZ}}(\varepsilon|p_{XY}) \subseteq \text{cl} \left( \bigcup_{m \geq 1} \underline{\mathcal{R}}_{\text{WZ}}(m, \varepsilon|p_{XY}) \right). \end{aligned} \quad (57)$$

Note here that  $\mathcal{R}_{\text{WZ}}(\varepsilon|p_{XY})$  is a closed set. Then from (57), we conclude that

$$\begin{aligned} \mathcal{R}_{\text{WZ}}(\varepsilon|W) &= \text{cl} \left( \bigcup_{m \geq 1} \underline{\mathcal{R}}_{\text{WZ}}(m, \varepsilon|p_{XY}) \right) \\ &= \text{cl} \left( \bigcup_{m \geq 1} \bigcap_{n \geq m} \mathcal{R}_{\text{WZ}}(n, \varepsilon|p_{XY}) \right), \end{aligned}$$

completing the proof.  $\blacksquare$

### B. Cardinality Bound on Auxiliary Random Variables

Define

$$\begin{aligned} \mathcal{Q}(p_{XY}) &\triangleq \{q_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|, q_{XY} = p_{XY}, \\ &\quad U \leftrightarrow X \leftrightarrow Y, X \leftrightarrow (U, Y) \leftrightarrow Z\}. \end{aligned}$$

We first prove the following lemma.

*Lemma 8:*

$$\begin{aligned} \underline{R}^{(\mu)}(p_{XY}) &\triangleq \min_{q \in \mathcal{Q}(p_{XY})} \{\bar{\mu} I_q(X; U|Y) + \mu E_q d(X, Z)\} \\ &= R^{(\mu)}(p_{XY}) \triangleq \min_{q \in \mathcal{P}_{\text{sh}}(p_{XY})} \{\bar{\mu} I_q(X; U|Y) + \mu E_q d(X, Z)\}. \end{aligned}$$

To prove Lemma 8, we set

$$\begin{aligned} \mathcal{Q}_1(p_{XY}) &\triangleq \{q_1 = q_{UXY} : |\mathcal{U}| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|, \\ &\quad q_{XY} = p_{XY}, U \leftrightarrow X \leftrightarrow Y\}, \end{aligned}$$

$$\begin{aligned} \mathcal{P}_1(p_{XY}) &\triangleq \{q_1 = q_{UXY} : |\mathcal{U}| \leq |\mathcal{X}|, \\ &\quad q_{XY} = p_{XY}, U \leftrightarrow X \leftrightarrow Y\}, \end{aligned}$$

$$\begin{aligned} \mathcal{Q}_2(q_{UXY}) &\triangleq \{q_2 = q_{Z|UXY} : \\ &\quad q_{UXYZ} = (q_{UXY}, q_2), X \leftrightarrow (U, Y) \leftrightarrow Z\}. \end{aligned}$$

By definition it is obvious that

$$\begin{aligned} \mathcal{Q}(p_{XY}) &= \{q = (q_1, q_2) : q_1 \in \mathcal{Q}_1(p_{XY}), \\ &\quad q_2 \in \mathcal{Q}_2(q_1)\}, \end{aligned} \quad (58)$$

$$\begin{aligned} \mathcal{P}_{\text{sh}}(p_{XY}) &= \{q = (q_1, q_2) : q_1 \in \mathcal{P}_1(p_{XY}), \\ &\quad q_2 \in \mathcal{Q}_2(q_1)\}. \end{aligned} \quad (59)$$

*Proof of Lemma 8:* We first observe that by (58), we have

$$\begin{aligned} &\underline{R}^{(\mu)}(p_{XY}) \\ &= \min_{q_1 \in \mathcal{Q}_1(p_{XY})} \min_{q_2 \in \mathcal{Q}_2(q_1)} \{\bar{\mu} I_{q_1}(X; U|Y) \\ &\quad + \mu E_{(q_1, q_2)} d(X, Z)\} \\ &= \min_{q_1 \in \mathcal{Q}_1(p_{XY})} \left\{ \bar{\mu} I_{q_1}(X; U|Y) \right. \\ &\quad \left. + \mu \min_{q_2 \in \mathcal{Q}_2(q_1)} E_{(q_1, q_2)} d(X, Z) \right\} \\ &= \min_{q_1 \in \mathcal{Q}_1(p_{XY})} \{\bar{\mu} I_{q_1}(X; U|Y) + \mu E_{(q_1, q_2^*(q_1))} d(X, Z)\}, \end{aligned}$$

where

$$q_2^* = q_2^*(q_1) = q_{Z|UY}^* = \{q_{Z|UY}^*(z|u, y)\}_{(u, y, z) \in \mathcal{U} \times \mathcal{Y} \times \mathcal{Z}}$$

is a conditional probability distribution that attains the following optimization problem:

$$\min_{q_2 \in \mathcal{Q}_2(q_1)} E_{(q_1, q_2)} d(X, Z).$$

We bound the cardinality  $|\mathcal{U}|$  of  $U$  to show that the bound  $|\mathcal{U}| \leq |\mathcal{X}|$  is sufficient to describe  $\underline{R}^{(\mu)}(p_{XY})$ . Observe that

$$p_X(x) = \sum_{u \in \mathcal{U}} q_U(u) q_{X|U}(x|u), \quad (60)$$

$$\begin{aligned} &\bar{\mu} I_{q_1}(X; U|Y) + \mu E_{(q_1, q_2^*)}(X, Z) \\ &= \sum_{u \in \mathcal{U}} p_U(u) \pi(q_{X|U}(\cdot|u)), \end{aligned} \quad (61)$$

where

$$\begin{aligned} &\pi(q_{X|U}(\cdot|u)) \\ &\triangleq \sum_{(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}} q_{X|U}(x|u) p_{Y|X}(y|x) q_{Z|UY}^*(z|u, y) \\ &\quad \times \log \left\{ \frac{q_{X|U}^{\bar{\mu}}(x|u)}{p_X^{\bar{\mu}}(x)} \frac{p_Y^{\bar{\mu}}(y) e^{-\mu d(x, z)}}{\left[ \sum_{\tilde{x} \in \mathcal{X}} p_{Y|X}(y|\tilde{x}) q_{X|U}(\tilde{x}|u) \right]^{\bar{\mu}}} \right\}. \end{aligned}$$

For each  $u \in \mathcal{U}$ ,  $\pi(q_{X|U}(\cdot|u))$  is a continuous function of  $q_{X|U}(\cdot|u)$ . Then by the support lemma,

$$|\mathcal{U}| \leq |\mathcal{X}| - 1 + 1 = |\mathcal{X}|$$

is sufficient to express  $|\mathcal{X}| - 1$  values of (60) and one value of (61). ■

Next we give a proof of Property 4 part a).

*Proof of Property 4 part a):* We bound the cardinality  $|\mathcal{U}|$  of  $U$  to show that the bound  $|\mathcal{U}| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$  is sufficient to describe  $\Omega(\alpha, \mu, \lambda)$  ( $p_{XY}$ ). Observe that

$$q_{XYZ}(x, y, z) = \sum_{u \in \mathcal{U}} q_U(u) q_{XYZ|U}(x, y, z|u), \quad (62)$$

$$\exp \left\{ -\Omega^{(\alpha, \mu, \lambda)}(q|p_{XY}) \right\} = \sum_{u \in \mathcal{U}} q_U(u) \Pi^{(\alpha, \mu, \lambda)}(q_{XYZ|U}(\cdot, \cdot, \cdot|u)), \quad (63)$$

where

$$\begin{aligned} & \Pi^{(\alpha, \mu, \lambda)}(q_{XYZ|U}(\cdot, \cdot, \cdot|u)) \\ & \triangleq \sum_{\substack{(x, y, z) \\ \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}} q_{XYZ|U}(x, y, z|u) \exp \left\{ -\lambda \omega_{q||p}^{(\alpha, \mu)}(x, y, z|u) \right\}. \end{aligned}$$

For each  $u \in \mathcal{U}$ ,  $\Pi^{(\alpha, \mu)}(q_{XYZ|U}(\cdot, \cdot, \cdot|u))$  is a continuous function of  $q_{XYZ|U}(\cdot, \cdot, \cdot|u)$ . Then by the support lemma,

$$|\mathcal{U}| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}| - 1 + 1 = |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$$

is sufficient to express  $|\mathcal{X}||\mathcal{Y}||\mathcal{Z}| - 1$  values of (62) and one value of (63). ■

### C. Proof of Property 2

In this appendix we prove Property 2. Property 2 part a) is a well known property. Proof of this property is omitted here. We only prove Property 2 part b).

*Proof of Property 2 part b):* Since  $\mathcal{P}^*(p_{XY}) \subseteq \mathcal{P}(p_{XY})$ , it is obvious that  $\mathcal{R}^*(p_{XY}) \subseteq \mathcal{R}(p_{XY})$ . Hence it suffices to prove that  $\mathcal{R}(p_{XY}) \subseteq \mathcal{R}^*(p_{XY})$ . We assume that  $(R, \Delta) \in \mathcal{R}(p_{XY})$ . Then there exists  $p \in \mathcal{P}(p_{XY})$  such that

$$R \geq I_p(U; X|Y) \text{ and } \Delta \geq E_p d(X, Z). \quad (64)$$

On the second inequality in (64), we have the following:

$$\begin{aligned} & \Delta \geq E_p d(X, Z) \\ & = \sum_{(u, y) \in \mathcal{U} \times \mathcal{Y}} p_{UY}(u, y) \left[ \sum_{z \in \mathcal{Z}} p_{Z|UY}(z|u, y) \right. \\ & \quad \times \left. \left( \sum_{x \in \mathcal{X}} d(x, z) p_{X|UY}(x|u, y) \right) \right] \\ & \geq \sum_{(u, y) \in \mathcal{U} \times \mathcal{Y}} p_{UY}(u, y) \\ & \quad \times \left[ \min_{z \in \mathcal{Z}} \left( \sum_{x \in \mathcal{X}} d(x, z) p_{X|UY}(x|u, y) \right) \right] \\ & = \sum_{(u, y) \in \mathcal{U} \times \mathcal{Y}} p_{UY}(u, y) \left( \sum_{x \in \mathcal{X}} d(x, z^*) p_{X|UY}(x|u, y) \right), \quad (65) \end{aligned}$$

where  $z^* = z^*(u, y)$  is one of the minimizers of the function

$$\sum_{x \in \mathcal{X}} d(x, z) p_{X|UY}(x|u, y).$$

Define  $\phi : \mathcal{U} \times \mathcal{Y} \rightarrow \mathcal{Z}$  by  $\phi(u, y) = z^*$ . We further define  $q = q_{UXYZ}$  by  $q_{UXY} = p_{UXY}$ ,  $q_Z = q_{\phi(U, Y)}$ . It is obvious that

$$q \in \mathcal{P}^*(p_{XY}) \text{ and } R \geq I_p(X; U|Y) = I_q(X; U|Y). \quad (66)$$

Furthermore from (65), we have

$$\Delta \geq E_p d(X, Z) \geq E_q d(X, \phi(U, Y)). \quad (67)$$

From (66) and (67), we have  $(R, \Delta) \in \mathcal{R}^*(p_{XY})$ . Thus  $\mathcal{R}(p_{XY}) \subseteq \mathcal{R}^*(p_{XY})$  is proved. ■

### D. Proof of Property 3

In this appendix we prove Property 3. From Property 2 part a), we have the following lemma.

*Lemma 9:* Suppose that  $(\hat{R}, \hat{\Delta})$  does not belong to  $\mathcal{R}(p_{XY})$ . Then there exist  $\epsilon, \mu^* > 0$  such that for any  $(R, \Delta) \in \mathcal{R}(p_{XY})$  we have

$$\bar{\mu}(R - \hat{R}) + \mu^*(\Delta - \hat{\Delta}) - \epsilon \geq 0.$$

Proof of this lemma is omitted here. Lemma 9 is equivalent to the fact that if the region  $\mathcal{R}(p_{XY})$  is a convex set, then for any point  $(\hat{R}, \hat{\Delta})$  outside the region  $\mathcal{R}(p_{XY})$ , there exists a line which separates the point  $(\hat{R}, \hat{\Delta})$  from the region  $\mathcal{R}(p_{XY})$ . Lemma 9 will be used to prove (5) in Property 3.

*Proof of (5) in Property 3:* We first recall the following definitions of  $\mathcal{P}(p_{XY})$  and  $\mathcal{P}_{\text{sh}}(p_{XY})$ :

$$\begin{aligned} \mathcal{P}(p_{XY}) & \triangleq \{p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}| + 1, U \leftrightarrow X \leftrightarrow Y, \\ & \quad X \leftrightarrow (U, Y) \leftrightarrow Z\}, \\ \mathcal{P}_{\text{sh}}(p_{XY}) & \triangleq \{p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}|, U \leftrightarrow X \leftrightarrow Y, \\ & \quad X \leftrightarrow (U, Y) \leftrightarrow Z\}. \end{aligned}$$

We prove  $\mathcal{R}_{\text{sh}}(p_{XY}) \subseteq \mathcal{R}(p_{XY})$ . We assume that  $(\hat{R}, \hat{\Delta}) \notin \mathcal{R}(p_{XY})$ . Then by Lemma 9, there exist  $\epsilon > 0$  and  $\mu^* > 0$  such that for any  $(R, \Delta) \in \mathcal{R}(p_{XY})$ , we have

$$\bar{\mu}^* \hat{R} + \mu^* \hat{\Delta} \leq \bar{\mu}^* R + \mu^* \Delta - \epsilon.$$

Hence we have

$$\begin{aligned} & \bar{\mu}^* \hat{R} + \mu^* \hat{\Delta} \leq \min_{(R, \Delta) \in \mathcal{R}(p_{XY})} \{\bar{\mu}^* R + \mu^* \Delta\} - \epsilon \\ & \stackrel{(a)}{=} \min_{p \in \mathcal{P}(p_{XY})} \{\bar{\mu}^* I_p(U; X|Y) + \mu^* E_p d(X, Z)\} - \epsilon \\ & \leq \min_{p \in \mathcal{P}_{\text{sh}}(p_{XY})} \{\bar{\mu}^* I_p(U; X|Y) + \mu^* E_p d(X, Z)\} - \epsilon \\ & = R^{(\mu^*)}(p_{XY}) - \epsilon. \quad (68) \end{aligned}$$

Step (a) follows from the definition of  $\mathcal{R}(p_{XY})$ . The inequality (68) implies that  $(\hat{R}, \hat{\Delta}) \notin \mathcal{R}_{\text{sh}}(p_{XY})$ . Thus  $\mathcal{R}_{\text{sh}}(p_{XY}) \subseteq \mathcal{R}(p_{XY})$  is concluded. We next prove  $\mathcal{R}(p_{XY}) \subseteq \mathcal{R}_{\text{sh}}(p_{XY})$ . We assume that  $(R, \Delta) \in \mathcal{R}(p_{XY})$ . Then there exists  $q \in \mathcal{P}(p_{XY})$  such that

$$R \geq I_q(X; U|Y), \Delta \geq E_q d(X, Z). \quad (69)$$

Then, for each  $\mu > 0$  and for  $(R, \Delta) \in \mathcal{R}(p_{XY})$ , we have the following chain of inequalities:

$$\begin{aligned} \bar{\mu}R + \mu\Delta &\stackrel{(a)}{\geq} \bar{\mu}I_q(X; U|Y) + \mu E_q d(X, Z) \\ &\geq \min_{q \in \mathcal{P}(p_{XY})} [\bar{\mu}I_q(X; U|Y) + \mu E_q d(X, Z)] \\ &= R^{(\alpha, \mu)}(p_{XY}). \end{aligned}$$

Step (a) follows from (69). Hence we have  $\mathcal{R}(p_{XY}) \subseteq \mathcal{R}_{\text{sh}}(p_{XY})$ .  $\blacksquare$

We next prove the two inequalities of (6) in Property 3.

*Proof of (6) in Property 3:* We first prove the second inequality of (6) in Property 3. We have the following chain of inequalities.

$$\begin{aligned} &R^{(\alpha, \mu)}(p_{XY}). \\ &\geq \min_{q \in \mathcal{P}(p_{XY})} [\bar{\mu}I_q(X; U|Y) + \mu E_q d(X, Z)] \\ &\stackrel{(a)}{=} \min_{q \in \mathcal{P}(p_{XY})} \left\{ \frac{\bar{\alpha}}{4\alpha} [D(q_X \| p_X) + D(q_{Y|XU} \| p_{Y|X} | q_{XU}) \right. \\ &\quad \left. + \gamma I_q(X; Z|UY)] \right. \\ &\quad \left. + \bar{\mu}I_q(X; U|Y) + \mu E_q d(X, Z) \right\} \\ &\geq \min_{q \in \mathcal{Q}} \left\{ \frac{\bar{\alpha}}{4\alpha} [D(q_X \| p_X) + D(q_{Y|XU} \| p_{Y|X} | q_{XU}) \right. \\ &\quad \left. + \gamma I_q(X; Z|UY)] \right. \\ &\quad \left. + \bar{\mu}I_q(X; U|Y) + \mu E_q d(X, Z) \right\} \\ &= \frac{1}{4\alpha} \tilde{R}^{(\alpha, \mu)}(p_{XY}). \end{aligned}$$

Step (a) follows from that when  $q \in \mathcal{P}(p_{XY})$ , we have

$$D(q_X \| p_X) = D(q_{Y|XU} \| p_{Y|X} | q_{XU}) = I_q(X; Z|UY) = 0,$$

We next prove the first inequality of (6) in Property 3. Let  $q_{\alpha, \mu}^* = q_{UXYZ, \alpha, \mu}^* \in \mathcal{Q}$  be a probability distribution which attains the minimum in the definition of  $\tilde{R}^{(\alpha, \mu)}(p_{XY})$ . Let  $\hat{q}_{\alpha, \mu} = \hat{q}_{UXYZ, \alpha, \mu}$  be a probability distribution with the form

$$\begin{aligned} &\hat{q}_{UXYZ, \alpha, \mu}(u, x, y, z) \\ &= q_{U|X, \alpha, \mu}^*(u|x) p_X(x) p_{Y|X}(y|x) q_{Z|UY, \alpha, \mu}^*(z|u, y). \end{aligned}$$

Define

$$\begin{aligned} \mathcal{Q}(p_{XY}) &\triangleq \{q_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|, q_{XY} = p_{XY}, \\ &\quad U \leftrightarrow X \leftrightarrow Y, X \leftrightarrow (U, Y) \leftrightarrow Z\}. \end{aligned}$$

By definition, we have  $\hat{q}_{\alpha, \mu} \in \mathcal{Q}(p_{XY})$ . Then we have the following chain of inequalities.

$$\begin{aligned} &\bar{\alpha} D(q_{\alpha, \mu}^* \| \hat{q}_{\alpha, \mu}) \\ &= \bar{\alpha} [D(q_{X, \alpha, \mu}^* \| p_X) + D(q_{Y|XU, \alpha, \mu}^* \| p_{Y|X} | q_{XU, \alpha, \mu}^*) \\ &\quad + I_{q_{\alpha, \mu}^*}(X; Z|UY)] \\ &\leq \bar{\alpha} [D(q_{X, \alpha, \mu}^* \| p_X) + D(q_{Y|XU, \alpha, \mu}^* \| p_{Y|X} | q_{XU, \alpha, \mu}^*) \\ &\quad + \bar{\mu} I_{q_{\alpha, \mu}^*}(X; Z|UY)] \\ &\quad + 4\alpha [\bar{\mu} I_{q_{\alpha, \mu}^*}(X; U|Y) + \mu E_{q_{\alpha, \mu}^*} d(X; Z)] \\ &= \tilde{R}^{(\alpha, \mu)}(p_{XY}) \end{aligned}$$

$$\begin{aligned} &\leq \bar{\alpha} [D(\hat{q}_{X, \alpha, \mu} \| p_X) + D(\hat{q}_{Y|XU, \alpha, \mu} \| p_{Y|X} | \hat{q}_{XU, \alpha, \mu}) \\ &\quad + I_{\hat{q}_{\alpha, \mu}}(X; Z|UY)] \\ &\quad + 4\alpha [\bar{\mu} I_{\hat{q}_{\alpha, \mu}}(X; U|Y) + \mu E_{\hat{q}_{\alpha, \mu}} d(X; Z)] \\ &= 4\alpha [\bar{\mu} I_{\hat{q}_{\alpha, \mu}}(X; U|Y) + \mu E_{\hat{q}_{\alpha, \mu}} d(X; Z)] \\ &\leq 4\alpha \log(|\mathcal{X}|e^{d_{\max}}). \end{aligned} \tag{70}$$

For simplicity of notation we set  $\xi \triangleq \log(|\mathcal{X}|e^{d_{\max}})$ . From (70) we have

$$D(q_{\alpha, \mu}^* \| \hat{q}_{\alpha, \mu}) \leq 4\xi \frac{\alpha}{\bar{\alpha}}. \tag{71}$$

By the Pinsker's inequality we have

$$\frac{1}{2} (\|q_{\alpha, \mu}^* - \hat{q}_{\alpha, \mu}\|_1)^2 \leq D(q_{\alpha, \mu}^* \| \hat{q}_{\alpha, \mu}) \tag{72}$$

From (71) and (72), we obtain

$$\|q_{\alpha, \mu}^* - \hat{q}_{\alpha, \mu}\|_1 \leq \sqrt{8\xi \frac{\alpha}{\bar{\alpha}}} \leq \frac{1}{2} \tag{73}$$

for any  $\alpha \in (0, \alpha_0]$  with

$$\alpha_0 = (32\xi + 1)^{-1} = [32 \log(|\mathcal{X}|e^{d_{\max}}) + 1]^{-1}. \tag{74}$$

The bound (73) implies that for any  $A \subseteq \{U, X, Y, Z\}$ , we have

$$\|q_{A, \alpha, \mu}^* - \hat{q}_{A, \alpha, \mu}\|_1 \leq \sqrt{8\xi \frac{\alpha}{\bar{\alpha}}} \leq \frac{1}{2}. \tag{75}$$

Then for any  $\alpha \in (0, \alpha_0]$ , we have the following chain of inequalities:

$$\begin{aligned} &|I_{q_{\alpha, \mu}^*}(U; X|Y) - I_{\hat{q}_{\alpha, \mu}}(U; X|Y)| \\ &\leq |H_{q_{\alpha, \mu}^*}(XY) - H_{\hat{q}_{\alpha, \mu}}(XY)| \\ &\quad + |H_{q_{\alpha, \mu}^*}(Y) - H_{\hat{q}_{\alpha, \mu}}(Y)| \\ &\quad + |H_{q_{\alpha, \mu}^*}(UXY) - H_{\hat{q}_{\alpha, \mu}}(UXY)| \\ &\quad + |H_{q_{\alpha, \mu}^*}(UY) - H_{\hat{q}_{\alpha, \mu}}(UY)| \\ &\stackrel{(a)}{\leq} -\sqrt{8\xi \frac{\alpha}{\bar{\alpha}}} \left[ \log \left( \sqrt{8\xi \frac{\alpha}{\bar{\alpha}}} \cdot \frac{1}{|\mathcal{X}||\mathcal{Y}|} \right) \right. \\ &\quad \left. + \log \left( \sqrt{8\xi \frac{\alpha}{\bar{\alpha}}} \cdot \frac{1}{|\mathcal{Y}|} \right) + \log \left( \sqrt{8\xi \frac{\alpha}{\bar{\alpha}}} \cdot \frac{1}{|\mathcal{U}||\mathcal{X}||\mathcal{Y}|} \right) \right. \\ &\quad \left. + \log \left( \sqrt{8\xi \frac{\alpha}{\bar{\alpha}}} \cdot \frac{1}{|\mathcal{U}||\mathcal{Y}|} \right) \right] \\ &= 2\sqrt{8\xi \frac{\alpha}{\bar{\alpha}}} \log \left\{ \left( \frac{\bar{\alpha}}{8\xi \alpha} \right) |\mathcal{U}||\mathcal{X}||\mathcal{Y}|^2 \right\} \\ &\stackrel{(b)}{\leq} 2\sqrt{8\xi \frac{\alpha}{\bar{\alpha}}} \log \left\{ \left( \frac{\bar{\alpha}}{8\xi \alpha} \right) |\mathcal{Z}||\mathcal{X}|^2|\mathcal{Y}|^3 \right\}. \end{aligned} \tag{76}$$

Step (a) follows from (75) and LEMMA 2.7 in Section 1.2 in Csiszár and Körner [4]. Step (b) follows from that when  $q \in \mathcal{Q}$ , we have  $|\mathcal{U}| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$ . On the other hand we have

$$|E_{q_{\alpha, \mu}^*} d(X, Z) - E_{\hat{q}_{\alpha, \mu}} d(X, Z)| \leq d_{\max} \sqrt{8\xi \frac{\alpha}{\bar{\alpha}}}. \tag{77}$$



Combining (76) and (77), we have

$$\begin{aligned}
& \left| \bar{\mu} I_{q_{\alpha,\mu}^*}(U; X|Y) + \mu E_{q_{\alpha,\mu}^*} d(X, Z) \right. \\
& \quad \left. - (\bar{\mu} I_{\hat{q}_{\alpha,\mu}}(U; X|Y) + \mu E_{\hat{q}_{\alpha,\mu}} d(X, Z)) \right| \\
& \leq |I_{q_{\alpha,\mu}^*}(U; X|Y) - I_{\hat{q}_{\alpha,\mu}}(U; X|Y)| \\
& \quad + |E_{q_{\alpha,\mu}^*} d(X, Z) - E_{\hat{q}_{\alpha,\mu}} d(X, Z)| \\
& \leq 2\sqrt{8\xi\frac{\alpha}{\bar{\alpha}}} \log \left\{ \left( \frac{\bar{\alpha}}{8\xi\alpha} \right) |\mathcal{Z}||\mathcal{X}|^2|\mathcal{Y}|^3 \right\} + d_{\max} \sqrt{8\xi\frac{\alpha}{\bar{\alpha}}} \\
& = c_1 \sqrt{\frac{\alpha}{\bar{\alpha}}} \log \left( c_2 \frac{\bar{\alpha}}{\alpha} \right). \tag{78}
\end{aligned}$$

Then we have

$$\begin{aligned}
& \frac{1}{4\alpha} \tilde{R}^{(\alpha,\mu)}(p_{XY}) \\
& = \frac{\bar{\alpha}}{4\alpha} [D(q_{X,\alpha,\mu}^* || p_X) + D(q_{Y|XU,\alpha,\mu}^* || p_{Y|X} | q_{XU,\alpha,\mu}^*) \\
& \quad + I_{q_{\alpha,\mu}^*}(X; Z|UY)] \\
& \quad + I_{q_{\alpha,\mu}^*}(X; U|Y) + \mu E_{q_{\alpha,\mu}^*} d(X; Z) \\
& \geq I_{q_{\alpha,\mu}^*}(X; U|Y) + \mu E_{q_{\alpha,\mu}^*} d(X; Z) \\
& \geq I_{\hat{q}_{\alpha,\mu}}(X; U|Y) + \mu E_{\hat{q}_{\alpha,\mu}} d(X; Z) \\
& \quad - c_1 \sqrt{\frac{\alpha}{\bar{\alpha}}} \log \left( c_2 \frac{\bar{\alpha}}{\alpha} \right) \\
& \stackrel{(a)}{\geq} \min_{q \in \mathcal{Q}(p_{XY})} \{I_q(X; Y|U) + \mu E_q d(X, Z)\} \\
& \quad - c_1 \sqrt{\frac{\alpha}{\bar{\alpha}}} \log \left( c_2 \frac{\bar{\alpha}}{\alpha} \right) \\
& \stackrel{(b)}{=} \min_{q \in \mathcal{P}_{\text{sh}}(p_{XY})} \{I_q(X; Y|U) + \mu E_q d(X, Z)\} \\
& \quad - c_1 \sqrt{\frac{\alpha}{\bar{\alpha}}} \log \left( c_2 \frac{\bar{\alpha}}{\alpha} \right) \\
& = R^{(\mu)}(p_{XY}) - c_1 \sqrt{\frac{\alpha}{\bar{\alpha}}} \log \left( c_2 \frac{\bar{\alpha}}{\alpha} \right). \tag{79}
\end{aligned}$$

Step (a) follows from that  $\hat{q}_{\alpha,\mu} \in \mathcal{Q}(p_{XY})$ . Step (b) follows from Lemma 8 stated in Appendix B. ■

#### E. Proof of Property 4 parts b), c), and d)

In this appendix we prove Property 4 parts b), c), and d).

*Proof of Property 4 parts b), c), and d):* We first prove parts b) and c). For simplicity of notations, set

$$\begin{aligned}
\underline{a} & \triangleq (u, x, y, z), \underline{A} \triangleq (U, X, Y, Z), \underline{\mathcal{A}} \triangleq \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \\
\omega_{q_{XYZ|U}}^{(\mu)}(x, y, z|u) & \triangleq g(\underline{a}), \\
\Omega^{(\alpha,\mu,\lambda)}(q|p_{XY}) & \triangleq \xi(\lambda).
\end{aligned}$$

Then we have

$$\xi(\lambda) = -\log \left[ \sum_{\underline{a} \in \underline{\mathcal{A}}} p_{\underline{A}}(\underline{a}) e^{-\lambda g(\underline{a})} \right].$$

We set

$$q^{(\lambda)}(u, x, y, z) \triangleq p_{\underline{A}}^{(\lambda)}(\underline{a}).$$

Then  $p_{\underline{A}}^{(\lambda)}(\underline{a}), \underline{a} \in \mathcal{A}$  has the following form:

$$p_{\underline{A}}^{(\lambda)}(\underline{a}) = e^{\xi(\lambda)} p_{\underline{A}}(\underline{a}) e^{-\lambda g(\underline{a})}.$$

By simple computations we have

$$\begin{aligned}
\xi'(\lambda) & = e^{\xi(\lambda)} \left[ \sum_{\underline{a} \in \underline{\mathcal{A}}} p_{\underline{A}}(\underline{a}) g(\underline{a}) e^{-\lambda g(\underline{a})} \right] \\
& = \sum_{\underline{a} \in \underline{\mathcal{A}}} p_{\underline{A}}^{(\lambda)}(\underline{a}) g(\underline{a}),
\end{aligned}$$

$$\xi''(\lambda) = -e^{2\xi(\lambda)}$$

$$\begin{aligned}
& \times \left[ \sum_{\underline{a}, \underline{b} \in \underline{\mathcal{A}}} p_{\underline{A}}(\underline{a}) p_{\underline{A}}(\underline{b}) \frac{\{g(\underline{a}) - g(\underline{b})\}^2}{2} e^{-\lambda\{g(\underline{a}) + g(\underline{b})\}} \right] \\
& = - \sum_{\underline{a}, \underline{b} \in \underline{\mathcal{A}}} p_{\underline{A}}^{(\lambda)}(\underline{a}) p_{\underline{A}}^{(\lambda)}(\underline{b}) \frac{\{g(\underline{a}) - g(\underline{b})\}^2}{2} \\
& = - \sum_{\underline{a} \in \underline{\mathcal{A}}} p_{\underline{A}}^{(\lambda)}(\underline{a}) g^2(\underline{a}) + \left[ \sum_{\underline{a} \in \underline{\mathcal{A}}} p_{\underline{A}}^{(\lambda)}(\underline{a}) g(\underline{a}) \right]^2.
\end{aligned}$$

By the Taylor expansion of  $\Omega^{(\alpha,\mu,\lambda)}(q|p_{XY})$  with respect to  $\lambda$  around  $\lambda = 0$ , we have

$$\begin{aligned}
& \Omega^{(\alpha,\mu,\lambda)}(q|p_{XY}) \\
& = \xi(\lambda) = \xi(0) + \xi'(0)\lambda + \frac{1}{2}\xi''(0)(\tau\lambda)^2 \\
& = E_q \left[ \omega_{q||p}^{(\alpha,\mu)}(X, Y, Z|U) \right] \\
& \quad - \frac{1}{2}(\tau\lambda)^2 \text{Var}_q \left[ \omega_{q||p}^{(\alpha,\mu)}(X, Y, Z|U) \right]
\end{aligned}$$

for some  $\tau \in [0, 1]$ . Then by the definition of  $\rho$ , we have

$$\begin{aligned}
& \Omega^{(\alpha,\mu,\lambda)}(q|p_{XY}) \\
& \geq \lambda E_q \left[ \omega_{q||p}^{(\alpha,\mu)}(X, Y, Z|U) \right] - \frac{1}{2}\rho\lambda^2.
\end{aligned}$$

The second inequality is obvious from the first inequality. We finally prove the part d). By the hyperplane expression  $\mathcal{R}_{\text{sh}}(p_{XY})$  of  $\mathcal{R}(p_{XY})$  stated Property 3 part b) we have that when  $(R + \tau, \Delta + \tau) \notin \mathcal{R}(p_{XY})$ , we have

$$\bar{\mu}^*(R + \tau) + \mu^*(\Delta + \tau) < R^{(\mu^*)}(p_{XY}) \tag{80}$$

for some  $\mu^* \in [0, 1]$ . Let  $\alpha_0 = \alpha_0(d_{\max}, |\mathcal{X}|)$  be the quantity defined by (7) in Property 3 part b). Then from (80), we have that for any  $\alpha \in (0, \alpha_0]$ ,

$$\begin{aligned}
& \bar{\mu}^* R + \mu^* \Delta + \tau \leq R^{(\mu^*)}(p_{XY}) \\
& \stackrel{(a)}{\leq} \frac{1}{4\alpha} \tilde{R}^{(\alpha,\mu^*)}(p_{XY}) + c_1 \sqrt{\frac{\alpha}{\bar{\alpha}}} \log \left( c_2 \frac{\bar{\alpha}}{\alpha} \right). \tag{81}
\end{aligned}$$

Step (a) follows from the first inequality of (6) in Property 3 part b). Fix any  $\delta > 0$ . We choose  $\alpha = \tau^{2+\delta}$ . For this choice of  $\alpha$ , we have (81) for any  $\tau \in (0, \alpha_0^{\frac{1}{2+\delta}}]$  and

$$\begin{aligned}
& c_1 \cdot \sqrt{\frac{\alpha}{\bar{\alpha}}} \log \left( c_2 \frac{\bar{\alpha}}{\alpha} \right) \times \frac{1}{\tau} \\
& = c_1 \frac{\tau^{\frac{\delta}{2}}}{\sqrt{1 - \tau^{2+\delta}}} \log \left( c_2 \cdot \frac{1 - \tau^{2+\delta}}{\tau^{2+\delta}} \right) \rightarrow 0 \text{ as } \tau \rightarrow 0.
\end{aligned}$$

Hence there exists a positive  $\nu = \nu(\delta, d_{\max}, |\mathcal{X}|, |\mathcal{Y}|, |\mathcal{Z}|)$  with  $\nu \leq \alpha_0^{\frac{1}{2+\delta}}$  such that for  $\tau \in (0, \nu] \subseteq (0, \alpha_0^{\frac{1}{2+\delta}}]$ ,

$$c_1 \cdot \sqrt{\frac{\alpha}{\bar{\alpha}}} \log \left( c_2 \frac{\bar{\alpha}}{\alpha} \right) \leq \frac{\tau}{2}. \quad (82)$$

The above inequality together with (81) yields that for  $\tau \in (0, \nu]$ ,

$$\bar{\mu}^* R + \mu^* \Delta + \frac{1}{2} \tau \leq \frac{1}{4\tau^{2+\delta}} \tilde{R}^{(\tau^{2+\delta}, \mu^*)}(p_{XY}). \quad (83)$$

Then for each  $\tau \in (0, \nu]$ , we have the following chain of inequalities:

$$\begin{aligned} & F(R, \Delta | p_{XY}) \\ & \geq \sup_{\lambda > 0} F^{(\tau^{2+\delta}, \mu^*, \lambda)}(\bar{\mu}^* R + \mu^* \Delta | p_{XY}) \\ & = \sup_{\lambda > 0} \frac{\Omega^{(\tau^{2+\delta}, \mu^*, \lambda)}(p_{XY}) - 4\lambda\tau^{2+\delta}(\bar{\mu}^* R + \mu^* \Delta)}{1 + 4(1 - \mu^* \tau^{2+\delta})\lambda} \\ & \stackrel{(a)}{\geq} \sup_{\lambda > 0} \frac{1}{1 + 4\lambda} \left\{ -\frac{1}{2}\rho\lambda^2 + \lambda\tilde{R}^{(\tau^{2+\delta}, \mu^*)}(p_{XY}) - 4\lambda\tau^{2+\delta}(\bar{\mu}^* R + \mu^* \Delta) \right\} \\ & \stackrel{(b)}{\geq} \sup_{\lambda > 0} \frac{1}{1 + 4\lambda} \left\{ -\frac{1}{2}\rho\lambda^2 + 2\tau^{3+\delta}\lambda \right\} \stackrel{(c)}{=} \frac{\rho}{2} g^2 \left( \frac{\tau^{3+\delta}}{\rho} \right). \end{aligned}$$

Step (a) follows from Property 4 part b). Step (b) follows from (83). Step (c) follows from an elementary computation. This completes the proof of Property 4 part d). ■

#### F. Proof of Lemma 1

To prove Lemma 1, we prepare a lemma. Set

$$\begin{aligned} \tilde{\mathcal{A}}_n & \triangleq \left\{ x^n : \frac{1}{n} \log \frac{p_{X^n}(x^n)}{q_{X^n}^{(i)}(x^n)} \geq -\eta \right\}, \\ \mathcal{A}_n & \triangleq \tilde{\mathcal{A}}_n \times \mathcal{M}_n \times \mathcal{Y}^n \times \mathcal{Z}^n, \\ \mathcal{A}_n^c & \triangleq \tilde{\mathcal{A}}_n^c \times \mathcal{M}_n \times \mathcal{Y}^n \times \mathcal{Z}^n, \\ \tilde{\mathcal{B}}_n & \triangleq \left\{ (s, x^n, y^n) : \frac{1}{n} \log \frac{p_{Y^n|X^n}(y^n|x^n)}{q_{Y^n|X^n S_n}^{(ii)}(y^n|x^n, s)} \geq -\eta \right\}, \\ \mathcal{B}_n & \triangleq \tilde{\mathcal{B}}_n \times \mathcal{Z}^n, \mathcal{B}_n^c \triangleq \tilde{\mathcal{B}}_n^c \times \mathcal{Z}^n, \\ \mathcal{C}_n & \triangleq \left\{ (s, x^n, y^n, z^n) : \right. \\ & \quad \left. \frac{1}{n} \log \frac{p_{X^n|S_n Y^n}(x^n|s, y^n)}{q_{X^n|S_n Y^n Z^n}^{(iii)}(x^n|s, y^n, z^n)} \geq -\eta \right\}, \\ \tilde{\mathcal{D}}_n & \triangleq \{(s, x^n, y^n) : s = \varphi^{(n)}(x^n), \\ & \quad q_{X^n|S_n Y^n}^{(iv)}(x^n|s, y^n) \leq M_n e^{n\eta} p_{X^n|Y^n}(x^n|y^n)\}, \\ \mathcal{D}_n & \triangleq \tilde{\mathcal{D}}_n \times \mathcal{Z}^n, \mathcal{D}_n^c \triangleq \tilde{\mathcal{D}}_n^c \times \mathcal{Z}^n. \end{aligned}$$

Then we have the following lemma.

*Lemma 10:*

$$\begin{aligned} p_{S_n X^n Y^n Z^n}(\mathcal{A}_n^c) & \leq e^{-n\eta}, p_{S_n X^n Y^n Z^n}(\mathcal{B}_n^c) \leq e^{-n\eta}, \\ p_{S_n X^n Y^n Z^n}(\mathcal{C}_n^c) & \leq e^{-n\eta}, p_{S_n X^n Y^n Z^n}(\mathcal{D}_n^c) \leq e^{-n\eta}. \end{aligned}$$

*Proof:* We first prove the first inequality. We have the following chain of inequalities:

$$\begin{aligned} p_{S_n X^n Y^n Z^n}(\mathcal{A}_n^c) & = p_{X^n}(\tilde{\mathcal{A}}_n^c) = \sum_{x^n \in \tilde{\mathcal{A}}_n^c} p_{X^n}(x^n) \\ & \stackrel{(a)}{\leq} \sum_{x^n \in \tilde{\mathcal{A}}_n^c} e^{-n\eta} q_{X^n}^{(i)}(x^n) \leq e^{-n\eta} \sum_{x^n} q_{X^n}^{(i)}(x^n) = e^{-n\eta}. \end{aligned}$$

Step (a) follows from the definition of  $\mathcal{A}_n$ . We next prove the second inequality. We have the following chain of inequalities:

$$\begin{aligned} p_{S_n X^n Y^n Z^n}(\mathcal{B}_n^c) & = p_{S_n X^n Y^n}(\tilde{\mathcal{B}}_n^c) \\ & \stackrel{(a)}{=} \sum_{(s, x^n, y^n) \in \tilde{\mathcal{B}}_n^c} p_{S_n X^n}(s, x^n) p_{Y^n|X^n}(y^n|x^n) \\ & \stackrel{(b)}{\leq} \sum_{(s, x^n, y^n) \in \tilde{\mathcal{B}}_n^c} e^{-n\eta} p_{S_n X^n}(s, x^n) q_{Y^n|S_n X^n}^{(ii)}(y^n|s, x^n) \\ & \leq e^{-n\eta} \sum_{s, x^n, y^n} p_{S_n X^n}(s, x^n) q_{Y^n|S_n X^n}^{(ii)}(y^n|s, x^n) = e^{-n\eta}. \end{aligned}$$

Step (a) follows from the Markov chain  $S_n \leftrightarrow X^n \leftrightarrow Y^n$ . Step (b) follows from the definition of  $\mathcal{B}_n$ . On the third inequality we have the following chain of inequalities:

$$\begin{aligned} p_{S_n X^n Y^n Z^n}(\mathcal{C}_n^c) & \stackrel{(a)}{=} \sum_{(s, x^n, y^n, z^n) \in \mathcal{C}_n^c} p_{X^n|S_n Y^n}(x^n|s, y^n) p_{S_n Y^n Z^n}(s, y^n, z^n) \\ & \stackrel{(b)}{\leq} \sum_{(s, x^n, y^n, z^n) \in \mathcal{C}_n^c} e^{-n\eta} q_{X^n|S_n Y^n Z^n}^{(iii)}(x^n|s, y^n, z^n) \\ & \quad \times p_{S_n Y^n Z^n}(s, y^n, z^n) \\ & \leq e^{-n\eta} \sum_{s, x^n, y^n, z^n} q_{X^n|S_n Y^n Z^n}^{(iii)}(x^n|s, y^n, z^n) \\ & \quad \times p_{S_n Y^n Z^n}(s, y^n, z^n) = e^{-n\eta}. \end{aligned}$$

Step (a) follows from the Markov chain  $X^n \leftrightarrow S_n Y^n \leftrightarrow Z^n$ . Step (b) follows from the definition of  $\mathcal{C}_n$ . We finally prove the fourth inequality. We have the following chain of inequalities:

$$\begin{aligned} p_{S_n X^n Y^n Z^n}(\mathcal{D}_n^c) & = p_{S_n X^n Y^n}(\tilde{\mathcal{D}}_n^c) \\ & = \sum_{s \in \mathcal{M}_n} \sum_{\substack{(x^n, y^n) : \varphi^{(n)}(x^n) = s, \\ p_{X^n|Y^n}(x^n|y^n) \leq (1/M_n) e^{-n\eta} \\ \times q_{X^n|S_n, Y^n}^{(iv)}(x^n|s, y^n)}} p_{X^n|Y^n}(x^n|y^n) p_{Y^n}(y^n) \\ & \leq \frac{e^{-n\eta}}{M_n} \sum_{s \in \mathcal{M}_n} \sum_{\substack{(x^n, y^n) : \varphi^{(n)}(x^n) = s, \\ p_{X^n|Y^n}(x^n|y^n) \leq (1/M_n) e^{-n\eta} \\ \times q_{X^n|S_n, Y^n}^{(iv)}(x^n|s, y^n)}} q_{X^n|S_n Y^n}^{(iv)}(x^n|s, y^n) p_{Y^n}(y^n) \\ & \leq \frac{e^{-n\eta}}{M_n} \sum_{s \in \mathcal{M}_n} \sum_{x^n, y^n} q_{X^n|S_n Y^n}^{(iv)}(x^n|s, y^n) p_{Y^n}(y^n) = e^{-n\eta}. \end{aligned}$$

*Proof of Lemma 1:* We set

$$\mathcal{E}_n \triangleq \left\{ (s, x^n, y^n, z^n) : \frac{1}{n} d(X^n, Z^n) \leq \Delta \right\}.$$

Set  $R^{(n)} \triangleq (1/n) \log M_n$ . By definition we have

$$\begin{aligned}
& p_{S_n X^n Y^n Z^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) \\
&= p_{S_n X^n Y^n Z^n} \left\{ \eta \geq \frac{1}{n} \log \frac{q_{X^n}^{(i)}(X^n)}{p_{X^n}(X^n)}, \right. \\
&\quad \eta \geq \frac{1}{n} \log \frac{q_{Y^n|X^n S}^{(ii)}(Y^n|X^n S)}{p_{Y^n|X^n}(Y^n|X^n)}, \\
&\quad \eta \geq \frac{1}{n} \log \frac{q_{X^n|S_n Y^n Z^n}^{(iii)}(X^n|S_n Y^n Z^n)}{p_{X^n|S_n Y^n}(X^n|S_n Y^n)}, \\
&\quad R^{(n)} + \eta \geq \frac{1}{n} \log \frac{q_{X^n|S_n Y^n}^{(iv)}(X^n|S_n Y^n)}{p_{X^n|Y^n}(X^n|Y^n)}, \\
&\quad \Delta \geq \frac{1}{n} \log \exp \{d(X^n, Z^n)\} \Big\}. \tag{84}
\end{aligned}$$

Then for any  $(\varphi^{(n)}, \psi^{(n)})$  satisfying

$$R^{(n)} = \frac{1}{n} \log M_n \leq R,$$

we have

$$\begin{aligned}
& p_{S_n X^n Y^n Z^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) \\
&\leq p_{S_n X^n Y^n Z^n} \left\{ \eta \geq \frac{1}{n} \log \frac{q_{X^n}^{(i)}(X^n)}{p_{X^n}(X^n)}, \right. \\
&\quad \eta \geq \frac{1}{n} \log \frac{q_{Y^n|X^n S}^{(ii)}(Y^n|X^n S)}{p_{Y^n|X^n}(Y^n|X^n)}, \\
&\quad \eta \geq \frac{1}{n} \log \frac{q_{X^n|S_n Y^n Z^n}^{(iii)}(X^n|S_n Y^n Z^n)}{p_{X^n|S_n Y^n}(X^n|S_n Y^n)}, \\
&\quad R + \eta \geq \frac{1}{n} \log \frac{q_{X^n|S_n Y^n}^{(iv)}(X^n|S_n Y^n)}{p_{X^n|Y^n}(X^n|Y^n)}, \\
&\quad \Delta \geq \frac{1}{n} \log \exp \{d(X^n, Z^n)\} \Big\}. \tag{85}
\end{aligned}$$

Hence, it suffices to show

$$\begin{aligned}
& P_c^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}; \Delta) \\
&\leq p_{S_n X^n Y^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) + 4e^{-n\eta}
\end{aligned}$$

to prove Lemma 1. By definition we have

$$P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) = p_{S_n X^n Y^n Z^n} (\mathcal{E}_n).$$

Then we have the following.

$$\begin{aligned}
& P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) = p_{S_n X^n Y^n Z^n} (\mathcal{E}_n) \\
&= p_{S_n X^n Y^n Z^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) \\
&\quad + p_{S_n X^n Y^n Z^n} ([\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n]^c \cap \mathcal{E}_n) \\
&\leq p_{S_n X^n Y^n Z^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) \\
&\quad + p_{S_n X^n Y^n Z^n} (\mathcal{A}_n^c) + p_{S_n X^n Y^n Z^n} (\mathcal{B}_n^c) \\
&\quad + p_{S_n X^n Y^n Z^n} (\mathcal{C}_n^c) + p_{S_n X^n Y^n Z^n} (\mathcal{D}_n^c) \\
&\stackrel{(a)}{\leq} p_{S_n X^n Y^n Z^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) + 4e^{-n\eta}.
\end{aligned}$$

Step (a) follows from Lemma 10.  $\blacksquare$

## G. Proof of Lemma 2

In this appendix we prove Lemma 2.

*Proof of Lemma 2:* We have the following chain of inequalities:

$$\begin{aligned}
& I(X_t; Y^{t-1} | S_n X_{t-1} Y_t^n) \\
&= H(Y^{t-1} | S_n X^{t-1} Y_t^n) - H(Y^{t-1} | S_n X^t Y_t^n) \\
&\leq H(Y^{t-1} | X^{t-1}) - H(Y^{t-1} | S_n X^n Y_t^n) \\
&\stackrel{(a)}{=} H(Y^{t-1} | X^{t-1}) - H(Y^{t-1} | X^n Y_t^n) \\
&\stackrel{(b)}{=} H(Y^{t-1} | X^{t-1}) - H(Y^{t-1} | X^{t-1}) = 0.
\end{aligned}$$

Step (a) follows from that  $S_n = \varphi^{(n)}(X^n)$  is a function of  $X^n$ . Step (b) follows from the memoryless property of the information source  $\{(X_t, Y_t)\}_{t=1}^\infty$ .  $\blacksquare$

## H. Proof of Lemma 6

In this appendix we prove (30) and (31) in Lemma 6.

*Proofs of (30) and (31) in Lemma 6:* By the definition of  $p_{X^t Z^t | S_n Y^n, q^t}(x^t, z^t | s, y^n)$ , for  $t = 1, 2, \dots, n$ , we have

$$\begin{aligned}
& p_{X^t Z^t | S_n Y^n, q^t}^{(\alpha, \mu, \theta; q^t)}(x^t, z^t | s, y^n) \\
&= C_t^{-1}(s, y^n) p_{X^t Z^t | S_n Y^n}(x^t, z^t | s, y^n) \\
&\quad \times \prod_{i=1}^t f_{p_i | q_i}^{(\alpha, \mu, \theta)}(x_i, y_i, z_i | u_i). \tag{86}
\end{aligned}$$

Then we have the following chain of equalities:

$$\begin{aligned}
& p_{X^t Z^t | S_n Y^n, q^t}^{(\alpha, \mu, \theta; q^t)}(x^t, z^t | s, y^n) \\
&\stackrel{(a)}{=} C_t^{-1}(s, y^n) p_{X^t Z^t | S_n Y^n}(x^t, z^t | s, y^n) \\
&\quad \times \prod_{i=1}^t f_{p_i | q_i}^{(\alpha, \mu, \theta)}(x_i, y_i, z_i | u_i) \\
&= C_t^{-1}(s, y^n) p_{X^{t-1} Z^{t-1} | S_n Y^n}(x^{t-1}, z^{t-1} | s, y^n) \\
&\quad \times \prod_{i=1}^{t-1} f_{p_i | q_i}^{(\alpha, \mu, \theta)}(x_i, y_i, z_i | u_i) \\
&\quad \times p_{X_t | Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) \\
&\quad \times f_{q_t | p_t}^{(\alpha, \mu, \theta)}(x_t, y_t | u_t) \\
&\stackrel{(b)}{=} \frac{C_{t-1}(s, y^n)}{C_t(s, y^n)} p_{X^{t-1} Z^{t-1} | S_n Y^n}^{(\alpha, \mu, \theta; q^{t-1})}(x^{t-1}, z^{t-1} | s, y^n) \\
&\quad \times p_{X_t | Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) \\
&\quad \times f_{q_t | p_t}^{(\alpha, \mu, \theta)}(x_t, y_t, z_t | u_t) \\
&= (\Phi_t^{(\alpha, \mu, \theta)}(s, y^n))^{-1} p_{X^{t-1} Z^{t-1} | S_n Y^n}^{(\alpha, \mu, \theta; q^{t-1})}(x_t, y_t, z_t | u_t) \\
&\quad \times p_{X_t | Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) \\
&\quad \times f_{q_t | p_t}^{(\alpha, \mu, \theta)}(x_t, y_t, z_t | u_t). \tag{87}
\end{aligned}$$

Steps (a) and (b) follow from (86). From (87), we have

$$\Phi_{t,q^t}^{(\alpha,\mu,\theta)}(s, y^n) p_{X^t Z^t | S_n Y^n}^{(\alpha,\mu,\theta)}(x^t, z^t | s, y^n) \quad (88)$$

$$\begin{aligned} &= p_{X^{t-1} Z^{t-1} | S_n Y^n}^{(\alpha,\mu,\theta;q^{t-1})}(x^{t-1}, z^{t-1} | s, y^n) \\ &\quad \times p_{X_t Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) \\ &\quad \times f_{q_t || p_t}^{(\alpha,\mu,\theta)}(x_t, y_t, z_t | u_t). \end{aligned} \quad (89)$$

Taking summations of (88) and (89) with respect to  $x^t, z^t$ , we obtain

$$\begin{aligned} &\Phi_{t,q^t}^{(\alpha,\mu,\theta)}(s, y^n) \\ &= \sum_{x^t, z^t} p_{X^{t-1} Z^{t-1} | S_n Y^n}^{(\alpha,\mu,\theta;q^{t-1})}(x^{t-1}, z^{t-1} | s, y^n) \\ &\quad \times p_{X_t Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) \\ &\quad \times f_{q_t || p_t}^{(\alpha,\mu,\theta)}(x_t, y_t, z_t | u_t), \end{aligned}$$

completing the proof.  $\blacksquare$

#### ACKNOWLEDGEMENT

I am very grateful to Dr. Shun Watanabe for his helpful comments.

#### REFERENCES

- [1] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 1-10, Jan. 1976.
- [2] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471-480, July 1973.
- [3] Y. Oohama and T. S. Han, "Universal coding for the Slepian-wolf data compression system and the strong converse theorem," *IEEE Transactions on Information Theory*, vol. IT-40, no.6, 1908-1919, Nov. 1994.
- [4] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. London, U.K., Academic, 1981.
- [5] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer-Verlag, Berlin, New York, 2002. The Japanese edition was published by Baifukan-publisher, Tokyo, 1998.
- [6] Y. Oohama, "Exponent function for one helper source coding problem at rates outside the rate region," *Proceedings of the 2015 IEEE International Symposium on Information Theory*, pp. 1575-1579, Hong Kong, China, June 14-19, 2015.
- [7] Y. Oohama, "Strong converse exponent for degraded broadcast channels at rates outside the capacity region," *Proceedings of the 2015 IEEE International Symposium on Information Theory*, pp. 939 - 943, Hong Kong, China, June 14-19, 2015.
- [8] Y. Oohama, "Strong converse theorems for degraded broadcast channels with feedback," *Proceedings of the 2015 IEEE International Symposium on Information Theory*, pp. 2510 - 2514, Hong Kong, China, June 14-19, 2015.